The Effects of Career and Technical Education: Evidence from the Connecticut Technical High School System

Eric Brunner, Shaun Dougherty, Stephen Ross

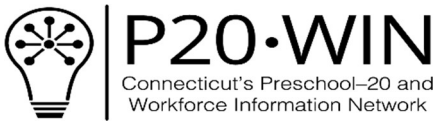
August 1, 2019

Abstract

We examine the effect of admission to 16 stand-alone technical high schools within the Connecticut Technical High School System (CTHSS) on student educational and labor market outcomes. To identify the causal effect of admission on student outcomes, we exploit the fact that CTHSS utilizes a score-based admissions system and identify the effect of admission using a regression discontinuity approach. We find that male students attending one of the technical high schools are approximately 10 percentage points more likely to graduate from high school and 8 percentage points less likely to attend college, although there is some evidence that the negative effects on college attendance fade over time. We also find that male students attending a technical high school have quarterly earnings that are approximately 31% higher. Analyses of potential mechanisms behind these results reveal that male students that attend a technical high school have higher 9th grade attendance rates and higher 10th grade test scores. We find little evidence that attending a technical high school affects the educational or labor outcomes of women. These effects appear relatively broad based across different types of students in that we find little evidence of heterogeneity in these effects over student attributes like race and ethnicity, free lunch eligibility or residence in a poor, central city school district. However, when distinguishing between students based on the Career and Technical Education (CTE) offerings of the high school that these students likely would have attended, we find that the effects of admission to a CTHSS school are noticeably larger when the counterfactual high school has less CTE offerings.

---

* Eric J. Brunner, Department of Public Policy, University of Connecticut, 10 Prospect Street, 4th Floor, Hartford, CT 06103, eric.brunner@uconn.edu; Shaun Dougherty, Department of Leadership, Policy, and Organizations, Vanderbilt University, 230 Appleton Place, Nashville, TN 37203-5721, shaun.dougherty@vanderbilt.edu; Stephen Ross, Department of Economics, University of Connecticut, 341 Mansfield Road, Unit 1063, Storrs, CT 06269-1063, stephen.l.ross@uconn.edu.



Acknowledgements: This research was made possible through data provided by the State of Connecticut's P20-WIN Program. We thank the Institute for Education Sciences for financial support under grant number R305A160195. We also thank Andrew Ju for excellent research assistance and for valuable comments Janet Currie, Caroline Hoxby, seminar participants at the Association for Public Policy Analysis and Management and the Society for Research on Educational Effectiveness for comments, and staff at the Performance Office of Connecticut State Department of Education.

I. Introduction

The past decade has witnessed a resurgence of interest in Career and Technical Education (CTE) as an alternative pathway for K-12 students. In 2015, more than 7 million secondary school students and nearly 4 million postsecondary students were enrolled in CTE programming (Passarella, 2018). Furthermore, in 2018, congress and the president reauthorized the Carl D. Perkins Career and Technical Education Act, providing $1.2 billion in funding for CTE programs and job training for students. As noted by Jacob (2017), proponents of CTE argue that such programming improves career opportunities by providing students with hands-on training and the soft skills necessary to succeed in the labor market. Similarly advocates stress that CTE may improve core academic skills by fostering student engagement and motivating students to attend school more regularly. Similarly, Cullen et al. (2013) in their recommendations for improving failing schools, argue for an increased focus on vocational training such as CTE as a way to foster practical skills and labor market integration. Nevertheless, despite the resurgence of interest in CTE, there is limited evidence on the effectiveness of CTE as an alternative pathway for K-12 students.

Carefully identified studies of CTE programs are limited to a small collection of studies, which in most cases, involve a small number of schools and/or students, and typically examine schools that volunteered to participate in an evaluation.[1] Kemple and Willner (2008) and Page (2012) examine labor market outcomes for approximately 1,400 students who applied for admission to nine different career academies that agreed to randomize admissions. Kemple and Willner (2008) find an 11% increase in earnings, but no effect on high school graduation, and Page (2012) in her follow-up shows that all effects are concentrated among males. Hemelt, Lenard and Paeplow (2019) exploit randomized admissions to

[1] A larger literature exploits longitudinal data to compare students in various CTE programs to observationally equivalent students in the same or similar schools or districts. Maxwell and Rubin (2002), Bishop and Main (2004), and Meer (2007) examine students surveyed in the National Educational Longitudinal Survey. Cellini (2006), Neumark and Rothstein (2006) and Kreisman and Stange (2017) examine students in the 1997 National Longitudinal Survey of Youth. Elliott, Hanser, and Gilroy (2002) and Maxwell and Rubin (2002) examine administrative student data from one or more school districts. On general, these studies tend to find positive effects on either educational attainment or earnings.

examine the effect of attending a single career academy in North Carolina and unlike Kemple and Willner (2008) find an improvement in graduation rates. Dougherty (2018) studies three Regional Vocational and Technical High Schools (RVTS) in Massachusetts that volunteered to be evaluated. Using a regression discontinuity strategy based on admission standards and outcomes for 4,000 applicants over three years, he finds a 7 to 10 percentage point increase in the likelihood of on-time graduation from high school. Cullen, Jacob and Levitt (2005) provide an exception to this pattern estimating the effects of 10 large career academies in Chicago Public schools using student residential proximity to schools as an instrument for school choice. They find large positive effects on graduation for students that attend career academies. However, since proximity predicts attendance at a career academy and also predicts the loss of students from the sending high schools, one cannot distinguish between the effect of attending the academy and the effects of losing potentially highly motivated students from the sending school.

In this study, we examine the effect of admission to any one of the 16 stand-alone technical high schools within the Connecticut Technical High School System (CTHSS).[2] We exploit the fact that CTHSS has a well-established, score-based admissions system and identify the effect of attendance using a regression discontinuity approach. Our analysis is based on the entire population of technical high schools in Connecticut and covers over 57,000 8th grade student applicants between 2006 and 2013. The roughly 11,000 students who attend these 16 schools each year comprise more than seven percent of all high school students in the state and all 16 of the technical high schools are oversubscribed and receive more applicants than they can accommodate. As a result, Connecticut presents a unique opportunity to examine the impact of stand-alone high school delivery of CTE where the services are being delivered at scale and a large fraction of the state's total student population is participating in the program. In fact, a key motivation behind Cullen et al.'s (2013) argument for expanding CTE options was potential concerns about scaling up and replicating successful models of school reform that emphasize college preparatory

[2] Our analysis excludes Bristol Technical because students at Bristol remain part of their original high school completing all or most of their academic coursework, as well as participating in both sporting and extracurricular activities, at their home school. The analysis also excludes Wright Technical, which did not open until 2014.

work, such as magnet schools. In this same vein, evidence on the effectiveness of CTE education at scale is very important given the magnitude of resources being devoted to CTE at both the federal and state level.

Students are admitted to CTHSS schools based on an application score. While this application score is not deterministic of admission, CTHSS school administrators establish a threshold each year prior to sending out initial letters of acceptance. Though the admission threshold is not made public nor observed in the admissions data, the pattern of admissions appears consistent with each school establishing such a threshold each year. We follow Porter and Yu (2015) and estimate an admissions threshold for each school and application year. We then use this threshold to identify the intent to treat and treatment on the treated effects of attending one of the CTHSS schools. First-stage regression discontinuity estimates show that students just above this threshold are 87% more likely to receive an acceptance letter. Similarly, male students are 58% more likely to attend the technical high school to which they applied as compared to students just below the threshold, and female students are 52% more likely. Given the strong gender patterns in selection into specific CTE programs, as well as traditional gender differences in labor market choices, we examine all outcomes using separate models for male and female applicants.[3]

Fuzzy regression discontinuity two-stage least squares estimates imply that male students attending one of the technical high schools are approximately 10 percentage points more likely to graduate from high school relative to the sample graduation rate of 83 percent. Male students are also 8 percentage points less likely to attend college relative to the sample college attendance rate of 40 percent. However, we also present evidence that suggests the negative effects on college attendance may fade over time. The higher high school graduation rates among male students are accompanied by a substantial

[3] For example in 2019, the programs of study on culinary arts, guest services, early child care and education, hairdressing and cosmetology, health technologies, hotel hospitality, and tourism captured 52 percent of female CTHSS students, but less than 7 percent of male CTHSS students; while the trade related programs of automotive manufacturing and technology, carpentry, collision repair, heavy equipment repair, electrical, HVAC, masonry, plumbing and welding contained 73 percent of all men and only 33 percent of women.

increase in earnings and labor market activity soon after leaving high school.[4] Total earnings increase by 37% relative to a sample average of $63,000, earnings per quarter increase 31% relative to an average of $6,000, and total number of quarters with earnings increases by 1.0 quarter relative to an average of almost 10 quarters. We find no significant effects of attending a CTHSS school for female students, and the point estimates are usually substantially smaller than the estimates for male students.

We undertake a series of analyses to examine the validity and robustness of these results. Our estimates are robust across different bandwidths, the use of a donut hole estimation strategy, and the inclusion of student attributes. The magnitude of the high school graduation effect is very stable as we decrease the bandwidth or add covariates. The college and labor market effects strengthen as the bandwidth is decreased, at worst consistent with an understatement of effects using our baseline bandwidth. Furthermore, we conduct balancing tests over applicant attributes for the overall sample and for the sample where we observe post-high school labor market outcomes within the State of Connecticut. For both samples and across different bandwidths student attributes appear to be balanced on either side of the admission threshold. We also conduct falsification tests by creating artificial thresholds above or below the real threshold, and find no evidence of differences in graduation rates, college attendance or labor market outcomes in these falsification tests. Finally, we must acknowledge that the distribution of application scores is not smooth and contains mass points and gaps where the density is well below the density at surrounding scores. This feature of our data precludes us from conducting a traditional McCrary tests for manipulation at the boundary. Consequently, in addition to the use of the donut hole strategy, we also present several simulation analyses to demonstrate that this irregular distribution appears to arise from the underlying components used to create the admission score, rather than manipulation at the boundary.

We next examine intermediate outcomes in high school as potential mechanisms for the increases in high school graduation. For male students, we find effects on 9th grade attendance and 10th grade test

[4] Depending on the cohort, we observe students in the labor market between one and a half and eight and a half years.

scores, but as before no effects for female students. We find approximately a two-percentage point increase in attendance rates over a base of 93 percent attendance. We also find that attending a technical high school increases average test scores by 18 percent of a standard deviation.

Finally, we examine heterogeneity in the estimated effects. We begin by examining whether treatment effects vary by eligibility for free and reduced-price lunch, race and ethnicity, and whether the applicant resided in one of the five central city districts in Connecticut. We find little evidence of heterogeneity across students over these dimensions. We then turn to identifying the counter-factual high schools for each student based on the high schools attended by individuals who were not admitted to the relevant CTHSS school and attended the same middle school as the student. We then measure the amount of CTE offerings at each counterfactual high school as a share of total electives offered at that school. We find evidence that the effect of attending a CTHSS school is negatively related to the extent of CTE offerings at the student's counterfactual high school. For example, a one standard deviation decrease in share of electives that are CTE offerings increases the return to attending a CTHSS school in terms of quarterly earnings by 12 percent. On the other hand, the share of electives that are CTE offerings in the area of human services and hospitality, which are fields dominated by female student enrollment, has no impact on the effect of CTHSS enrollment on student outcomes.

II. Connecticut Technical High School System

The Connecticut Technical High School System (CTHSS) is a quasi-independent school district of choice comprised of 16 high schools where all students who attend the school participate in some form of CTE. Specifically, all students take CTE coursework in lieu of other electives. CTE coursework is often grouped into one of 10 to 17 programs of study, such as information technology, health services, cosmetology, heating ventilation and air conditioning, or production processes. Within these programs, students will take at least three (often more) aligned courses, all focused on this program of study. Often, these sequences will also be combined with career awareness activities and opportunities for related work-based learning in settings outside of the school building and day. In contrast, traditional comprehensive high schools offer at most only 2 to 4 CTE programs from which to choose. While some

schools offer multi-course sequences similar to the CTHSS schools, often students may choose to take only one or two courses, sometimes not even in the same program. At CTHSS, 9th grade students explore 3 to 6 different possible programs of interest and at the end of that year prioritize which programs they would consider pursuing for the rest of high school. Beginning in grade 10, they are assigned a program based on preferences and availability. They then spend the next three years with a stable cohort of peers and instructors within their CTE coursework. Further, instructors from the CTE courses often collaborate with teachers in core academic areas to ensure overlap of content.

The schools serve students across all of Connecticut and are located throughout the state. The roughly 11,000 students who attend these 16 schools comprise more than seven percent of all high school students in the state, implying that the system provides services at scale to a substantial share of Connecticut's high school student population. Importantly, 31 percent of total enrollment comes from the state's five poorest central city school districts of Bridgeport, Hartford, New Haven, New London and Waterbury, and in aggregate the CTHSS serves a disproportionate number of students from lower-income families. The breadth of the students and environments served by this district offer a unique opportunity to understand not only how such an educational delivery model might impact student outcomes, but also how these effects might differ by setting or student characteristics.

Eighth graders across the state can elect to apply in the winter before they would enroll in 9th grade to attend high school at one of the CTHSS schools.[5] Students can apply to multiple schools and must rank-order their choices. All 16 of the technical high schools are oversubscribed and receive more applicants than they can accommodate. Admission to schools is coordinated by the central office of the CTHSS district. Students send a copy of their application to their first-choice school, the first-choice school processes the application and issues acceptance letters, and the processed information including subcomponent scores is placed into an applicant database for all schools and stored at the CTHSS district office. Each student receives an application score following a common standardized formula. For the 9th

[5] A limited number of 9th grade students apply for admission to a CTHSS school for 10th grade. These students are dropped from our analysis sample.

grade years of 2006-07 through 2008-09, the score is based on standardized 8th grade test scores in math and language arts (reading and writing) plus GPA and attendance in middle school. For the 9th grade years of 2009-10 through 2013-14, two additional categories were added based on points for extracurricular activities and a written statement.[6] The number of points associated with each component in each application year is shown in Table 1. The first-choice school is responsible for scoring the application and all schools use that score. Even though the underlying attendance and standardized test scores are close to continuous, the scoring system discretizes each of these components into an ordinal set of points that are then added together to form the score. The discrete nature of these components when combined with the high correlation between them yields a distribution of raw scores that is irregular with both mass points and holes/gaps in what might otherwise appear as a smooth distribution.

Within each school and year, school administrators have described establishing an admissions threshold and then sending out initial acceptance letters primarily to students whose scores lie above the threshold. Later waves of letters are sent out possibly based on lower thresholds until all seats in the school are filled. However, there exists some deviation between the student's score and whether the student receives an acceptance letter. Some students may be admitted with lower scores in order to increase diversity. Other students with higher scores may not be admitted because they applied late, were below an initial threshold for admission and withdrew their application prior to a second wave of admissions, or were excluded based on information in their disciplinary file. The recording of acceptance letters is also imperfect given we observe some students in the CTHSS system the year after they applied even though there is no record of them receiving an acceptance letter. Therefore, the admissions process results in a "fuzzy" discontinuity where the noise arises from deviations of school administrators from the scoring system, errors in the recording of acceptance letters, and imperfect take-up by applicants upon receipt of an acceptance letter. Based on inspection, the discontinuity appears to arise around the initial wave of acceptance letters (especially in terms of school attendance since being in the initial wave has a

[6] Points for extracurricular activities and the written statement are based on information provided by the applicant.

large effect on take-up). Finally, special education applicants, i.e. applicants with an Individualized Educational Plan or IEP, are subject to another layer of review and evaluation after their score is calculated, and so enter another layer of review prior to admission. As a result, we exclude students with an IEP from our analysis.

III. Methods

We model the relationship between our outcomes and the admission score using a "fuzzy" regression discontinuity design (Imbens & Lemeiux, 2008; Murnane & Willett, 2011). We employ a local-linear regression framework using a uniform kernel. Unlike most such analyses, however, we do not observe the threshold established for sending out the initial waves of application letters. Therefore, we identify the thresholds empirically for each school and year following Porter and Yu's (2015) recommendation that the thresholds/discontinuity be estimated by selecting the threshold that yields the largest discontinuity in the dependent variable. Specifically, we estimate linear probability models for receiving an acceptance letter (T_{isyt}) separately for each school *s* and application year *y* for the sample of applicants *i* from town *t* controlling for linear running variables in the admissions score (X_{isyt}) on either side of candidate thresholds or cut-offs (X^*_{sy}):

$$T_{isyt} = \alpha_{sy} d\left(X^*_{sy} \leq X_{isyt}\right) + \theta_{11}\left(X_{isyt} - X^*_{sy}\right) + \theta_{12}\left(X_{isyt} - X^*_{sy}\right) d\left(X^*_{sy} \leq X_{ist}\right) + \varepsilon_{1ist} \qquad (1)$$

where $d\left(X^*_{sy} \leq X_{isyt}\right)$ is a binary indicator that takes the value of one if the condition is satisfied. Equation (1) is estimated using observations that fall within bandwidth BW or for which:

$X_{isyt} \in [X^*_{sy} - BW, X^*_{sy} + BW]$,

and the threshold estimate is selected as:

$\widehat{X^*_{sy}} = argmax_{X^*_{sy}}\, \widehat{\alpha_{sy}}(X^*_{sy})$ over all $X^*_{sy} \in [X_{min} + BW, X_{max} - BW]$

For more details, please see Section 1 of the methodological appendix.

We then create a centered score, $\tilde{X}_{isyt} = X_{isyt} - \widehat{X^*_{sy}}$ and pool the data across schools and years in order to estimate models of student outcomes (y) using 2SLS:

$$y_{isyt} = \beta A_{isyt} + \theta_{21}\tilde{X}_{isyt} + \theta_{22} X_{isyt} d\left(0 \leq \tilde{X}_{isyt}\right) + \delta_{2sy} + \gamma_{2t} + \varepsilon_{2ist} \qquad (2)$$

where A_{isyt} denotes whether the individual attends the technical high school to which they applied in the following year, δ_{2sy} is a vector of school-by-application year fixed effects, and γ_{2t} is a vector of applicant town of residence fixed effects effectively identifying the likely counterfactual high school or schools.

Finally, the first stage equation for A_{isyt} is:

$$A_{isyt} = \tilde{\alpha}\, d\left(0 \leq \tilde{X}_{isyt}\right) + \theta_{31}\tilde{X}_{isyt} + \theta_{32}X_{isyt}d\left(0 \leq \tilde{X}_{isyt}\right) + \delta_{3sy} + \gamma_{3t} + \varepsilon_{3is} \quad (3)$$

where $\tilde{\alpha}$ represents the composite or sample average effect of being above the threshold on treatment.

The parameter of interest is β in equation (2), the population causal effect of participating in a CTHSS school for students who are just above the standardized admissions threshold compared to those who just missed the threshold. As is standard in regression discontinuity analyses, our estimates represent the Local Average Treatment Effect (LATE) for students near the admissions threshold for each school and year (see Angrist, Imbens and Rubin, 1996), but in our case the variation in cut-offs across schools and over time yields estimates that are representative over a broader set of student scores. Standard errors are clustered two ways following our fixed effects structure: application school by application year and sending town.[7]

As will be shown in more detail later in the paper, the empirical distribution of raw admission scores contains a large number of mass points that lie above what might otherwise be a smooth unimodal distribution. Further, the centered score distribution suggests that cut-offs tend to fall at these mass points, potentially because the admission process proceeds downwards through the distribution until a desired number of applicants is reached and such a process will be more likely to stop on a mass point than on any other specific point in the distribution. This feature of the data makes it impossible to conduct traditional McCrary tests for manipulation at the boundary. Therefore, in order to address concerns about

[7] Many prior studies with discrete running variables have clustered their standard errors by the running variable. However, as demonstrated by Kolesár & Rothe (2018) clustering by the running variable leads to confidence intervals that have poor coverage properties. Furthermore, as an alternative to our two-way clustering, we also conduct inference using finite-sample exact randomization inference tests following Cattaneo et al. (2019). The resulting *p*-values are substantially smaller than the ones presenting in the paper.

bias from manipulation, we estimate models using a donut hole approach restricting the sample by dropping observations at the cut-off for the school and year of those observations, or more specifically observations are selected into the sample based on: $X_{isyt} \in \left[X^*_{sy} - BW, X^*_{sy} - 1\right]$ or $\left[X^*_{sy} + 1, X^*_{sy} + BW\right]$ (Barreca et al., 2011; Barreca et al., 2016; Card & Giuliano, 2014; Canaan & Mouganie, 2018).

IV. Data and Sample

Our sample is comprised of the approximately 57,000 8th graders who apply to one of the 16 technical high schools for the academic years running from 2006-07 to 2013-14. The sample contains one observation for every application submitted even if a student submits multiple applications. The first stage take-up rate in the fuzzy RD will adjust for the fact that students can attend at most one of these schools.[8] The CTHSS admissions data contains each student applicant's name, date of birth, home town, middle school, the total admissions score, the individual components of the score, and in most cases the State Assigned Student Identification Number (SASID). We match the CTHSS admissions records to the Connecticut State Department of Education's (CSDE) longitudinal data system sequentially using the following criteria: SASID, exact match on first and last name plus birth year, first initial and exact match on last name plus birth year and month, and exact match on last name plus exact birth date. The reason for the sequential match process is reporting errors in the CTHSS application file on birth dates, spelling errors and uses of nicknames in the application file that parents and/or students fill out by hand. A failure to match after applying all of these criteria leads to the observation being omitted from the sample. Our resulting match rate was 95 percent yielding a final sample of 57,658 student applications.

From the CSDE longitudinal data system we obtained information on each student's race, gender, free or reduce price lunch status, English learner and special education status, i.e. presence of an IEP. The CSDE data also provides information on short- and medium-term educational outcomes including:

[8] By clustering on the student's town of residence, we assure that all applications submitted by any student are by construction included in the same cluster.

standardized test scores prior to and during high school, attendance, high school graduation, as well as information on college attendance drawn from the National Student Clearinghouse. Further, through the P20Win process established for the State of Connecticut, students in our sample are matched to Connecticut State Department of Labor (CSDOL) data. This CSDOL match is facilitated by Department of Motor Vehicle records that contain gender, birth date, and first and last name and is matched to the CSDOL data using social security numbers. CSDOL personnel then match the resulting data to the CSDE data using an exact match on birth date and gender and a fuzzy match algorithm on name. The fuzzy match confidence requires an estimated confidence of 70%, which yields a match rate of 72.3% between the student applicant records and the CSDOL data.[9] Failure to match applicants in the CSDOL data may be driven by several factors including never having a driver's license in the state of Connecticut, name changes due to marriage or other factors, moving out of state prior to or upon the completion of high school or failure to participate in the labor market after high school perhaps due to college attendance. Our labor market data ends in the 1st quarter of 2018. Therefore, we restrict this sample to the years 2006 to 2012 so that we have a potential for at least 6 quarters of data on each applicant

We selected a primary bandwidth of 10, and then test the sensitivity of our results to the use smaller bandwidths where there is less opportunity for bias.[10] Table 2 is intended to demonstrate the generalizability of our analysis by presenting sample means for the state overall, for the applicant pool and for the portion of the applicant pool within our baseline bandwidth of 10 points on either side of the school specific cut-off. The first panel presents means for student attributes and standardized test scores, which are available for the state overall. The second panel presents means for the CTHSS application score and the key score components including standardized tests, grades and attendance, which are only available for applicants. The CTHSS applicant sample is substantially less female (42%) than the student

[9] A fuzzy match criteria of 60% only yields an additional 500 matches, many of which looked erroneous upon visual inspection by CSDOL personnel.

[10] Our optimal bandwidth calculations follow Calonico, Cattaneo and Farrell (2018). Such methods are not ideal for discrete running variables so as noted above we use that bandwidth as a starting point and then test for robustness by reducing the bandwidth.

population statewide (49%). African-American, Hispanic and Free-lunch eligible students are substantially over represented relative to state wide averages with percent African-American being 50 percent higher and percent Hispanic and Free-lunch eligible almost double the shares statewide, and standardized test scores are about 2/3rd of a standard deviation below the state-wide averages. The means around the estimated admissions cut-offs are quite similar on student demographics to the means of the within bandwidth sample, but standardized test scores and admissions components are modestly lower for the within bandwidth sample, differences are always less than 1/2 a standard deviation. More importantly, however, the standard deviations of the variables do not fall very much as we move from the full sample of applicants to the sample within 10 points of the cut-off. The standard deviation of the standardized test scores only falls by between 15 and 20 percent, and the standard deviation of the admissions test score components which are directly truncated only falls by 43%. The modest reduction in variances within the regression discontinuity sample arises in large part because of the variance in cut-offs across schools and application years.

As mentioned above, we split the sample between male and female students both because they tend to pursue very different programs in the CTHSS system (for example the most common program among female students is cosmetology) and because historically men and women have different patterns of labor force participation. Next, we drop both students who apply in 9th grade for admission to the CTHSS for 10th grade and special education/IEP students given that the CTHSS system treats both of these groups quite differently in the admissions process. Finally, we also drop students who we cannot find in the CSDE data base in 9th grade because we have no outcomes for those students. The resulting sample includes 25,072 male applications and 20,983 female applications after dropping the 3,245 male and the 1,912 female 10th grade applicants, the 6,644 male applications and 2,328 female applications who are special education students and the 1,148 male and 1,083 female applicants who are not observed in 9th grade.

In Table 3, we present the means of outcome variables and student attributes for the full sample, the sample within 10 points of the cut-off, the sample within 10 points above the cut-off, and the sample

within 10 points below the cut-off. The first four columns present these means for male students and the last four columns present the same means for female students. Females applicants are substantially more likely to attend college and have lower earnings. Female applicants are also more likely to be minorities, free lunch eligible and English language learners. As before, the means for the full sample and the sample within the bandwidth are quite similar except that test scores are modestly lower in the bandwidth sample. When we compare the samples above and below the cut-offs, the means on socio-demographic attributes are also quite similar, but we observe substantial differences in educational attainment, test scores and labor market outcomes. Naturally, these variables correlate with the admissions score, but the formal regression discontinuity analysis will control for that correlation using a running variable. We also formally test for, and demonstrate, balance among the predetermined demographic characteristics later in the paper.

Modelling the First Stage Regression Discontinuity

As discussed above, the CTHSS system does not establish and document specific binding thresholds for school admission each year, and so we empirically select a threshold for each school and application year. We estimate equation (1) separately for each school and year identifying the cut-off score that maximizes the discontinuity in the probability of receiving an acceptance letter.[11] We then estimate a first stage equation by pooling the data from all schools and years and imposing a donut hole specification by dropping observations at the selected threshold for each school and year.[12] Figure 1a and Table 4 columns 1 and 2 present the pooled estimates for whether a student receives an acceptance letter using our standard 10-point bandwidth. Figures 1b and 1c and the additional columns of Table 4 present first-stage estimates by gender for whether we observe the student in the specific school to which they applied in the next academic year. All figures show a clear discontinuity with the probability of receiving

[11] The sending of an acceptance letter is recorded in the system by the date on which the acceptance letter was sent. Students are also coded as having received an acceptance letter if the system records a date at which the student responded to and accepted the offer, even if no date is recorded for the sending of the acceptance letter.

[12] We do not impose the donut hole when selecting the thresholds because the dropping of sample at each threshold could create non-convexities in the function that we are maximizing in selecting the optimal threshold.

an acceptance letter being above 0.9 and approaching one as the running variable increases past the cut-off. Figures 2b and 2c show a different pattern with attendance falling with the running variable to the right of the cut-off, consistent with higher scoring students having more options or coming from better school districts on average and thus being more likely to turn down the offer. As shown in Table 4, the estimated first stage effect of being above the cut-off on receiving an acceptance letter is 0.86 implying a large 86 percentage point increase in the likelihood of receiving a letter. The first stage for being observed in the technical high school is somewhat smaller, but still very sizable, at 0.58 for men and 0.52 for women. The results are presented in Table 4 in pairs of columns with and without controls and the estimates are virtually identical. As shown in Table 1A of the Empirical Appendix, we find similar first stage estimates for the labor market sample.

The use of the sample for selecting the threshold does not affect inference in the second stage of the 2SLS models because those models simply require that the instrument have sufficient power and that the exclusion restriction be satisfied along with the standard correction for using a constructed prediction of the endogenous regressor. However, the OLS/clustered standard errors in the first stage model may be biased because the thresholds were selected using the same sample, and the dummy variable based on that estimated threshold is included as a regressor. If this bias is severe, our use of *F*-statistics from the first stage regression to evaluate the strength of the instrument may be misleading. In order to address this concern, we follow Card, Mas and Rothstein (2008) and draw a holdout sample using the non-holdout sample for each school and year to estimate the threshold. We use the pooled holdout samples to estimate the first stage regression and repeat the entire exercise for four different draws of the holdout sample. The resulting *F*-statistics are always very strong, ranging between 456 and 628 for the full sample and between 458 and 674 for the donut hole sample. We also estimate the first stage models separately for each school and year using the hold out samples. The means of the estimated discontinuities over all schools and years range between 0.525 and 0.540 for the four simulations, respectively, and the fraction of thresholds that are significant at the 10 percent level ranges between 0.795 and 0.843. Further, the average discontinuity at the threshold across years for any school or simulation never falls below 0.329,

and the fraction of significant thresholds for any school in a simulation never falls below 0.571. Therefore, the empirically selected thresholds together provide a very strong instrument for explaining school attendance broadly across the entirely sample of CTHSS schools. Admittedly, the standard errors on the reduced form student outcome estimates presented later in the paper suffer from a similar problem, but these results are simply presented for completeness and we primarily rely on our 2SLS estimates for inference. For more details, see Section 1.1. of the Methodological Appendix.

Distribution of the Running Variable

One of the key assumptions of any RD analysis is that individuals cannot precisely control the value of the running variable and so cannot adjust that variable strategically in response to the cut-off value (Lee and Lemieux, 2010). In the current context, it is highly unlikely that students could manipulate their position along the running variable and strategically sort to the right side of the cut-off, given the admission process used by CTHSS. Specifically, recall that schools set the threshold after observing the pool of applicants, all their scores and all their provided materials. As a result, it is nearly impossible for students to know the exact position of the cut-off, which varies across schools and years. On the other hand, there is potentially more room for manipulation by the school personnel that are charged with processing and calculating each applicant's total score. However, manipulation also seems unlikely in this situation. First, school administrators have no incentive to manipulate the score because they are free to depart from the threshold that they have set for sending out letters and face little oversight or scrutiny in terms of their individual admissions decisions. Further, recall that the four primary components that make up a student's total score are standardized 8th grade test scores in math and language arts (reading and writing) plus GPA and attendance in middle school. All of these components are clear and objective with little to no room for manipulation. There is potentially more room for the manipulation of the final two components of the total score, namely the points assigned for extracurricular activities and a student's written statement for why they want to attend a technical high school, but those components do not enter the decision process until 2009. Again however, even this strategy seems unlikely, given that students applying to the technical schools most likely never interacted with the school officials at the technical

high schools. We also verify that the components sum to the composite score for all students in our sample so any manipulation would have to be sophisticated with school administrators manipulating a component in order to get a desired result on the total score.

In order to illustrate the issues that arise in the score distribution and prevent the use of a standard McCrary test, we plot the raw score distribution separately for 06-08, 09-10 and 11-13 since 2009 is when the points for extracurricular activities and the written statement were added and in 2011 the number of points assigned to those criteria was increased from 6 to 20. Note that the scores in 2009-2013 are above 100 because the maximum score was increased when the extra two components were added. The left hand side of Figure 2 shows these unconditional score distributions. Notice that the distributions exhibit substantial mass points and some gaps or holes in the distribution relative to any smooth distribution one might fit to this data. Further, in 06-08 when the distribution is based entirely on objective information, there is a large cliff in the unconditional distribution at around 45. In the bottom half of the figure, we repeat these raw distributions except that we drop students whose scores are exactly equal to the cut-off for their school and application year (donut hole distribution). If the irregularity of the distribution was driven primarily by manipulation, we would expect that most of the mass points in the distribution for 06-08 should have disappeared when we dropped students at these cut-offs, but the general nature of the score distribution appears unchanged by this sample restriction. However, the cliff in 06-08 does disappear when we drop students at the cut-off providing motivation for estimating our models using a donut hole specification.

We also conduct a series of simulations to ensure that the irregular features of the score distribution arise from the processes that generate the distribution, rather than manipulation around the boundary. The first two simulations focus on the raw score distribution. We first simulate the cut-offs for each school and year drawing from the distribution of observed cut-offs in each year for the other schools in order see how much of the deviation caused by dropping students at cut-offs that arose in the data and tended to occur at mass points, as opposed to being explained explicitly by the location of the cut-off for that school in that year. Specifically, we calculate the root mean squared error arising from comparing the

left and right hand side of the distributions shown in Figure 2 separately for each subsample, and then calculate a similar root mean squared error where we draw the cut-off randomly based on selecting from the cut-offs chosen by other schools in the same year. The deviation arising from drawing simulated cut-offs from a smoothed distribution of scores explains 71, 62 and 60 percent of the deviation that arises from using the cut-off itself for the periods of 06-08, 09-10 and 11-13, respectively. In our opinion, this correlation is substantial given that the application pools differ substantially across schools likely leading to differences in mass points across schools.

We next simulate raw score distributions year-by-year by measuring the correlation between the score components and randomly drawing components from the empirical distribution of those components in a manner that preserves this correlation. This allows us to create populations of fake total scores that are not influenced by manipulation around specific thresholds. We then plot the distributions of these simulated scores. While this simulation cannot perfectly replicate the raw distribution, the simulations do consistently generate similar shaped distributions for each year with significant numbers of both mass points and gaps.

Finally, the centered score distributions sometimes contain what looks like cliffs where the density is high above the cut-off and then drops substantially immediately below the cut-off (much like the cliff observed in the raw distribution for 2006-08). We speculate that the process of admitting applicants until a quota has been met could cause the cut-off to fall on or just after mass points. For each school, we draw a random number of students to receive admissions letters from the empirical distribution of the number of students admitted each year over all years. We then use this simulated number of admissions to create a simulated cut-off score. For 2006 through 2009, the simulated centered score distributions match the actual distribution quite well, including replicating cliffs observed in the data. However, beginning in 2010, the simulated distribution becomes much smoother and the empirical distribution continues to have significant cliffs. Therefore, we replicate all results in the paper dropping

the data from 2010 and beyond, and the results are robust.[13] The results of the simulations are shown in the methodological appendix, and the results of the pre-2010 subsample regressions are shown in the Empirical Appendix Table 2A. All of the simulations above are described, and the detailed results are presented, in Section 2 of the Methodological Appendix.

Balancing Tests

To further rule out manipulation at the boundary, we conduct balancing tests intended to detect sorting on observables across the cut-off boundaries using our donut hole sample. For both the male and female samples pooled across years and schools, we regress a dummy variable for whether the applicant's score is above the cut-off on student attributes including race and ethnicity, whether the student is free lunch eligible and whether the student is an English language learner plus the linear running variable for the student's score.[14] Columns 1-3 of Table 5 presents these results for bandwidths of 6, 8 and 10 points on either side of the cut-off for the sample of male students, while columns 4-6 present the same results for the sample of the female students. None of the individual variables are significant, and the *F*-tests on the set of variables is far from significant. Table 6 repeats this analysis for the sample of male and female students that are matched to the Connecticut State Department of Labor data. Our sample continues to pass balance at the cut-off, and we do not observe any substantial effect of attending a CTHSS school and being present in the labor market sample. Appendix Table 3A presents the same balancing tests for the full sample without dropping observations at the thresholds. Passing the balancing tests with the full sample provides further support of our maintained assumption that the non-standard distribution arises from the data, as opposed to manipulation.

We conduct the balancing tests with the cut-off on the left-hand side so that we can avoid type I error by conducting a single *F*-test for whether the student attributes are unrelated to the cut-off after

[13] The one exception is college attendance. In our main results we find that male students that attend a technical high school are approximately 8 percentage points less likely to attend college. However, when we restrict the sample to 2006-2009 application years, we find that male students that attend a technical high school are no less likely to attend college (see appendix Table 2A). We discuss this in more detail in the results section of the paper.

[14] As with our main RD models, these balancing tests include school by application year fixed effects and applicant town of residence fixed effects.

conditioning on the running variable. However, this specification requires a constant slope for the running variable because we cannot condition on the cut-off since the cut-off is the dependent variable. Therefore, we re-estimate the balancing tests one student attribute at a time with the student attributes on the left-hand side, the dummy for the cut-off on the right-hand side and controls for the running variable that allow the slope to differ above and below the cut-off. Again, parameter estimates on the cut-off dummy are insignificant for all student attributes, samples and bandwidths. These results are shown in Appendix Tables 4A and 5A for the donut hole and full samples, respectively.

V. Results

In this section we present our core findings concerning the effect of attending a CTHSS school on educational attainment and labor market outcomes. We begin by presenting outcomes related to educational attainment and then turn to labor market outcomes. Figure 3 presents a traditional reduced form regression discontinuity graph with a centered score, fitted line to the running variable on either side of zero, and the mean of the outcome at each score on the vertical axis. The top half of the figure presents results for male students, and the bottom half presents results for female students. High school graduation rates are shown on the left hand side, and college attendance rates are shown on the right side. For male students, we observe a classic pattern with the score means forming a relatively tight scatterplot around the fitted lines with a clear discontinuity implying higher rates of high school graduation and lower rates of college attendance for applicants above the threshold. In contrast, for female students, there is little evidence of a discontinuity in either high school graduation or college attendance rates at the cut-off. Figure 4 presents similar results for both quarterly earnings and number of quarters with earnings for the labor market sample. Again, for the sample of male students only, we observe a clear discontinuity at the threshold for both quarterly earnings and the number of quarters with earnings, with both labor market outcomes increasing just above the threshold.

Table 7 presents the reduced form estimates of being above the admission threshold at one of the CTHSS schools on high school graduation, attendance at any college, attendance at a two-year college and attendance at a four-year college using the donut hole sample. The first panel presents the results for

male students while the second panel presents results for female students. The estimates are presented in pairs of columns with the first column showing results from models that omit the controls from the balancing tests and the second column presenting results based on including those controls. As with Figure 3, we find large effects for all four outcomes in the male sample where students just above the cut-off are 6 percentage points more likely to graduate from high school and 4.5 percentage points less likely to attend any college with the college effects being divided relatively evenly between 2- and 4-year colleges. The estimates with controls are very similar (differ by less than 2%) to the estimates without controls. Importantly however, our finding that attending a technical high school reduces male college attendance holds only in the full sample when we combine all application years. As noted previously, in Table 2A of the Empirical Appendix, we replicate these reduced form results dropping the data from application years 2010 forward. As shown in Table 2A, when we restrict the sample to the early application years (2006-2009), we find no evidence that attending a technical high school reduces male college attendance. This finding may represent evidence that the effects on college attendance fade over time, which would be consistent with the results of Bifulco et al. (2014) who find that the effects of high school peers on college attendance tend to fade over time.[15] As shown in second panel, for the sample of female students, the point estimates are very close to zero and statistically insignificant. The estimates using the full sample are qualitatively similar and shown in Table 6A.[16]

Table 8 presents the same set of results for total earnings, quarterly earnings and number of quarters with earnings.[17] Being above the cut-off implies a 21 percent increase in total earnings, a 19 percent increase in quarterly earnings, an increase in the number of quarters worked of over one-half of

[15] However, we can not rule out other explanations for this result until we obtain additional waves of data on college attendance which will allow us to examine whether the negative effects on college attendance fade over time for the entire sample.

[16] The estimates for high school graduation for the full sample are somewhat larger and the estimates for attending college are somewhat smaller, but the estimates differ by only about 10 percent. The estimates also continue to be very stable to the addition of controls.

[17] As noted above, the labor market regressions are run with a smaller sample of individuals who were matched to the Connecticut State Department of Labor data. Table 7A reports estimates for the educational attainment outcomes for this subsample, and the estimates are very similar to the results in Table 7.

one quarter and an increase in quarterly earnings measured at age 23 of 20 percent. Once again, the inclusion of controls or the use of the full sample (Appendix Table 8A) has little impact on the estimates.[18] As with educational attainment, we estimate relatively precise zeros for female applicants.

Table 9 presents the two stage least squares estimates for the sample of male students allowing us to assess the impact of attending one of the technical high schools and verifying the statistical significance of the effect given that the 2SLS standard errors are not affected by the use of estimated thresholds. The resulting treatment on the treated effects are relatively large. In the top panel (full sample of male students), attending one of the technical high schools increases high school graduation rates by 10 percentage points relative to a sample average of 83%, and reduces attendance at any college by 8 percentage points relative to a 40% average. In the labor market (second panel), attending a technical high school results in 37 percent higher total earnings, 31 percent higher quarterly earnings and 32 percent higher quarterly earnings at age 23. For comparison, the average quarterly earnings in the sample of male students was approximately $6,000. Attending a technical high school also results in 1.0 additional quarters with earnings relative to a sample average of approximately 10 post-high school quarters with earnings. While large, these labor market effects are reasonable given the likely impact of high school graduation on wages and the increase in labor supply made possible by the reduction in college enrollment. We do not present the 2SLS results for women because the estimates are all close to zero as shown by the reduced form. The 2SLS estimates for the full sample of men (no donut hole) is shown in Table 9A and as before the estimates are quite similar for the two samples.[19]

Potential Mechanisms – In School Outcomes

[18] In this case, the estimates for total and quarterly earnings are very similar between the donut hole and full samples, and the estimates for number of quarters of earnings and quarterly earnings at 23 are smaller in the full sample. Therefore, we do not observe any systematic evidence of smaller estimates in the donut hole sample, which might potentially arise from the elimination of manipulation bias.

[19] As noted earlier, the running variable is discrete raising concerns about traditionally clustered standard errors. As an alternative, we conduct finite-sample exact randomization inference tests for the rejection of the null hypothesis of no discontinuity at the threshold following Cattaneo et al. (2019). The resulting *p*-values are shown in Appendix Table 10A and suggest substantially more confidence in our estimates than the precision implied by the two way (school by year and town) clustered standard errors presented in the paper.

In this section we examine potential mechanisms for the improved rates of high school graduation and level of success in the labor market and whether there is evidence of heterogeneous treatment effects.[20] Columns 1 - 4 of Table 10 present reduced-form and 2SLS estimates of the effect of attending a CTHSS school on 9th grade attendance and standardized 10th grade test scores, average of reading and mathematics. Focusing on the 2SLS estimates, we find that attendance rates improve by 1.7 percentage points relative to a sample mean rate of 93 percent. Test scores improve by 18 percent of a standard deviation. The estimates suggest that both increased attendance and improved mastery of academic course content could play a role in the higher rates of high school graduation and labor market success.

Additional Robustness Tests

In Table 11, we present reduced form estimates for our main outcomes based on alternative bandwidths. The first two panels present results based on 6-point and 8-point bandwidths, while the third panels present results from Tables 7 and 8 based on baseline bandwidth of 10-points. The effects on high school graduation are robust and stable across the bandwidths with the estimates only differing from the baseline estimate by no more than 5 percent. However, the effects for college attendance and labor market strengthen for smaller bandwidths. For total, quarterly and quarterly at age 23 earnings, the increases are moderate ranging between 13 and 28 percent for the 10-point and 6-point bandwidths. The estimated effects are most sensitive to bandwidth choice for college attendance and number of quarters worked with approximately a 75 percent increase in the magnitude of estimates between the 10-point and 6-point bandwidths. We have verified that all estimates for the sample of female students remain small and insignificant when we reduce the bandwidth to 6 points.

In general, results should typically be viewed as robust if they are stable or increase as the bandwidth is narrowed. Optimal bandwidths are selected as a trade-off between precision that is gained as the bandwidth is expanded and the possible bias that arises from a bandwidth that includes points further

[20] For consistency, we estimate 9th grade attendance for the same sample where we observe high school graduation, but for the test score sample we must omit students in the applications years of 2012 and 2013 because the State of Connecticut stopped statewide testing of 10th graders in order to pilot in select school districts the use of the SAT test as the 10th grade standardized exam.

away from the discontinuity (Calonico, Cattaneo and Farrell, 2018). Shrinking the bandwidth should reduce bias at the cost of some loss in the precision of the estimates. For all our educational and labor market outcomes the 6-point bandwidth yields robust and large effects. However, note that given our discrete running variable, as we reduce the bandwidth the estimated slope of the running variable is based on a smaller and smaller number of discrete points. This potentially raises the concern that the estimated slope of the running variable could be heavily influenced by one or two points. Therefore, in order to be conservative, we continue to focus our attention on the 10-point bandwidth estimates.

In Table 12, we present falsification tests where we move the cut-off down 10 points or up 10, 15 or 20 points with the 10-point bandwidth so that the true discontinuity is excluded from the sample.[21] We present reduced form estimates because the first stage has no power at the false cut-offs (see Appendix Table 11A). Panels 1-4 of Table 11 show the estimates for different falsification tests starting with the false cut-off being set 10 points below the true cut-off and moving to 20-points above the true cut-off. In all four panels, the point estimates are substantially smaller than the reduced form estimates in Tables 7 and 8, and with one exception statistically insignificant. Furthermore, in the one case where the estimate is significant at the 10 percent level (Panel 4, column 6) it is the opposite sign of the actual estimate reported in Table 8.

Heterogeneous Treatment Effects

In Table 13, we examine whether the effects of attending a technical high school are heterogenous across different students based on free lunch eligibility (Panel 1), being either African-American or Hispanic (Panel 2), or residence in one of Connecticut's five poorest central cities (Panel 3).[22] We estimate models in which we interact treatment, the running variables and the school by cohort fixed effects with the student attribute and focus on the 2SLS models in order to conduct inference on the interactions, where we use being above the threshold and its interaction with the attribute as instruments.

[21] We cannot move the threshold down by more than 10 points because the distribution of centered scores becomes very thin and the resulting falsification test would have very little power.
[22] The five cities included in this list are Bridgeport, Hartford, New Haven, New London and Waterbury.

Columns 1-4 present the estimates for graduation, college attendance, quarterly earnings, and number of quarters with earnings, respectively. The vast majority of the interaction estimates are insignificant. Only the interaction of central city residence with treatment in the model for number of quarters with earnings is significant, and this result is significant at the 99 percent confidence level. Therefore, even with the large number of tests for heterogeneity, it appears that most of the labor supply effects on the extensive margin are concentrated among students from one of Connecticut's poorest central cities.

We next examine whether the treatment effect varies by the availability of Career and Technical Education (CTE) in the high school the student would have attended if they had not attended a CTHSS school. In order to determine this counterfactual high school or high schools, we organize students into cells of applicants by the town or city in which they resided in 8th grade when applying to a CTHSS school.[23] For each of these cells, we record the high schools attended by individuals who were not admitted to the CTHSS school to which they applied. We then use State of Connecticut administrative files to calculate the number of CTE courses offered in each non-CTHSS school as a share of the total number of electives offered by that school, and average these shares across the students' counterfactual high schools.[24] Table 14 presents the means and standard deviations on the share of electives that are CTE overall, trade focused CTE, and CTE focused on human services, tourism or hospitality (services).[25] The average share of CTE courses overall is twice as larger in CTHSS schools as compared to our counterfactual high schools. When focusing on trades, these differences are magnified with on average non-CTHSS schools offering only about 5% of their electives as trade focused courses and trade offerings representing over half of the courses in CTHSS schools. On the other hand, in terms of human services

[23] The focus on towns is in large part justified because school district boundaries in Connecticut follow town and city boundaries except where smaller towns have been consolidated to form a combined regional high school district.

[24] In most towns, the share variables are primarily driven by the CTE offerings for the town high school or high schools because most students attend the town high schools. However, in some of the larger cities, the counter-factual is an average across a combination of large traditional city high schools and many magnet high schools located in or around the city.

[25] For trade focused, we select all courses in the state data based in the categories of manufacturing, transportation, and architecture which primarily contains programs related to building trades; while for the services category we select all courses in human services, tourism/hospitality, and family and consumer sciences.

and hospitality the average CTHSS high school actually offers a somewhat smaller share of electives in these areas compared to the counterfactual high schools.

In Table 15, we present models where we interact the share of elective courses at a student's counterfactual high school that are CTE with treatment. Panel 1 presents the results using all CTE offerings, Panel 2 presents the results using the offering of CTE courses related to the trades, and Panel 3 presents results related to Human Services and Tourism. We have relatively consistent results across graduation, college attendance and quarterly earnings with declining effects on each of these as the counterfactual high school offers an increasing share of electives as CTE course or as trade focused CTE courses. These relatively consistent findings stand in contrast to the earlier heterogeneity analysis that showed little evidence of heterogeneity. In contrast, for CTE courses focused on human services or tourism, there is no evidence of any heterogeneous effects: all of the point estimates on the interaction terms are relatively small in magnitude and statistically insignificant. Given the concentration of women in the human services and tourism CTE areas, these results are consistent with our findings of no treatment effects for women from attending a CTHSS school.

For comparison purposes, we examine how much of the gap between CTHSS and non-CTHSS schools is eliminated by a standard deviation increase in either CTE offerings overall or CTE trade offerings at the counterfactual high school. Looking back to Table 14, a one standard deviation increase in share of CTE courses at non-CTHSS schools is 0.067, which is 15.3 percent of the gap between CTE offerings at CTHSS and non-CTHSS schools, and so eliminating 15.3 percent of that gap in course offerings is associated with reducing the effects on college attendance by 6.5 percent and reducing the effects on quarterly earnings by 3.4 percent. Similarly, focusing on trade offerings within CTE, a one standard deviation increase is 0.023, which is only 4.5 percent of the gap in trade offerings between CTHSS and non-CTHSS schools. In comparison, the standardized effects of admission to a CTHSS school on high school graduation, college attendance and quarterly earnings are 1.7, 1.9 and 1.3 percent, which are sizable given the small reduction in the trade focused opportunity gap between CTHSS and non-CTHSS high schools.

If we use the reduction in the gap in offerings between non-CTHSS and CTHSS schools to scale the estimated effects, we find that closing 10 percent of the gap reduces the estimated positive treatment effects on quarterly earnings for all CTE courses by 2.2 percent and graduation and quarterly earnings for trades by 3.7 and 3.0 percent, respectively. Therefore, while a significant part of the gain from attending a CTHSS school can be attributed solely to the increase in CTHSS offerings, these findings leave substantial room for other features of the CTHSS experience to affect outcomes. These larger effects must arise from various aspects of the integrated CTE experience offered in the CTHSS system, rather than simply access to more CTE offerings.[26]

VI. Discussion

In this study, we examine the effect of attending one of Connecticut's 16 stand-alone technical high schools on student educational and labor market outcomes. We find large, robust positive effects for male students on high school graduation and post-high school labor market outcomes with the improved labor market outcomes being accompanied by declines in college enrollment. These positive outcomes may in part be due to the positive effects of CTHSS attendance on in-school outcomes such as attendance and 10th grade test scores. These estimates are robust to alternative bandwidths, the inclusion of controls for student attributes and a donut hole specification. We also demonstrate that the sample exhibits balance on student socio-demographic attributes. Falsification tests cannot identify similar discontinuities at false thresholds that are set above or below the true cut-offs, further bolstering the validity of the fuzzy regression discontinuity design.

We also examine whether these impacts differ by the CTE offerings that a student likely would have had available if they had not attended a CTHSS school. Specifically, we identify a counterfactual high school based on students who attended the same middle school and applied to the same CTHSS

[26] A second explanation could be that the return to CTE courses is non-linear and rises as more CTE offerings are provided. In those circumstances, the marginal effects of a change in traditional CTE offers would be below the effects created by the larger share of courses in CTHSS schools. In order to test this later hypothesis, we re-estimated models incorporating the interaction of the square of the CTE variable with treatment again using the interaction with being above the threshold as an instrument, but we found evidence of decreasing returns to additional CTE offerings as the number of offerings increase.

school, but were not admitted. We then measure the number of CTE offerings in each counterfactual school as a share of total elective course offerings at that high school. We find that students who likely would have attended a school with minimal CTE offerings benefit more in terms of quarterly earnings from attending one of the CTE schools. When focusing on CTE offerings in the trades, we find that less trade offerings at the counterfactual high school increases the effects of attending CTE for both high school graduation and quarterly earnings. We show that the influence of counterfactual offerings on returns to attending CTHSS schools is sizable by scaling these reductions by the reduction in the gap in offerings between non-CTHSS and CTHSS schools and find that closing 10 percent of the gap reduces the estimated positive treatment effects by 2 to 4 percent. Therefore, part of the gain from attending a CTHSS school can be attributed solely to the increase in CTHSS offerings. However, these findings leave substantial room for returns arising from the integrated provision of CTE in a standalone setting as provided by the CTHSS in the State of Connecticut.

Our study offers some of the first quasi-experimental evidence of the treatment effects associated with a technical high school program that is being offered at scale and involves the evaluation of all schools in the program, not just schools that volunteer. Career and technical education has been an important strategy for improving the economic opportunities of students who might not pursue a traditional four year college degree, and this study documents large positive effects for male students who participate in these programs. The evidence of the effectiveness of CTE, especially in terms of labor force participation as captured by quarters with earnings, is especially important given the declining opportunities for non-college educated workers and the declining labor force participation among non-college going prime age males (Abraham and Kearney 2018; Aguiar, Bils, Charles, and Hurst 2017; Autor 2019; Austin, Glaeser, and Summers 2018).

References

Abraham, K. G., & Kearney, M. S. (2018). Explaining the decline in the US employment-to-population ratio: A review of the evidence (No. w24333). National Bureau of Economic Research.

Aguiar, M., Bils, M., Charles, K. K., & Hurst, E. (2017). Leisure luxuries and the labor supply of young men (No. w23552). National Bureau of Economic Research.

Almond, D., Doyle Jr, J. J., Kowalski, A. E., & Williams, H. (2010). Estimating marginal returns to medical care: Evidence from at-risk newborns. *The Quarterly Journal of Economics*, 125(2), 591-634.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434), 444-455.

Austin, B. A., Glaeser, E. L., & Summers, L. H. (2018). Jobs for the Heartland: Place-based policies in 21st century America (No. w24548). National Bureau of Economic Research.

Autor, D. (2019). Work of the past, work of the future (No. w25588). National Bureau of Economic Research.

Barreca, A. I., Guldi, M., Lindo, J. M., & Waddell, G. R. (2011). Saving babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics*, 126(4), 2117-2123.

Barreca, A. I., Lindo, J. M., & Waddell, G. R. (2016). Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1), 268-293.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. *The Quarterly Journal of Economics*, 119(1), 249-275.

Bifulco, R., Fletcher, J. M., Oh, S. J., & Ross, S. L. (2014). Do high school peers have persistent effects on college attainment and other life outcomes?. *Labour Economics*, 29, 83-90.

Bishop, J. H., & Mane, F. (2004). The impacts of career-technical education on high school labor market success. *Economics of Education Review*, 23(4), 381-402.

Canaan, S., & Mouganie, P. (2018). Returns to education quality for low-skilled students: Evidence from a discontinuity. *Journal of Labor Economics*, 36(2), 395-436.

Card, D., & Giuliano, L. (2014). Does gifted education work? For which students? (No. w20453). National Bureau of Economic Research.

Card, D., Mas, A., & Rothstein, J. (2008). Tipping and the Dynamics of Segregation. *The Quarterly Journal of Economics*, 123(1), 177-218.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2018). Optimal bandwidth choice for robust bias corrected inference in regression discontinuity designs. Working Paper.

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2019). The Regression Discontinuity Design. In Handbook of Research Methods in Political Science and International Relations, eds. L. Curini and R. J. Franzese, Sage Publications.

Cellini, S. R. (2006). Smoothing the transition to college? The effect of Tech-Prep programs on educational attainment. *Economics of Education Review*, 25(4), 394-411.

Cullen, J. B., Jacob, B. A., & Levitt, S. D. (2005). The impact of school choice on student outcomes: an analysis of the Chicago Public Schools. *Journal of Public Economics*, 89(5-6), 729-760.

Cullen, J. B., Levitt, S. D., Robertson, E., & Sadoff, S. (2013). What can be done to improve struggling high schools?. *Journal of Economic Perspectives*, 27(2), 133-52.

Department of Education, Office of Vocational and Adult Education, Investing in America's Future: A Blueprint for Transforming Career and Technical Education, Washington, D.C., 2012.

Dougherty, S. M. (2018). The effect of career and technical education on human capital accumulation: Causal evidence from Massachusetts. *Education Finance and Policy*, 13(2), 119-148.

Elliott, M. N., Hanser, L. M., & Gilroy, C. L. (2002). Career academies: Additional evidence of positive student outcomes. *Journal of Education for Students Placed At Risk*, 7(1), 71-90.

Hemelt, S. W., Lenard, M. A., & Paeplow, C. G. (2019). Building bridges to life after high school: Contemporary career academies and student outcomes. *Economics of Education Review*, 68, 161-178.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615-635.

Kemple, J. J., & Willner, C. J. (2008). Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood (pp. 4-5). New York, NY: MDRC.

Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108(8), 2277-2304.

Kreisman, D., & Stange, K. (2017). Vocational and Career Tech Education in American High Schools: The Value of Depth Over Breadth. *Education Finance and Policy*, 1-72.

Maxwell, N. L., & Rubin, V. (2002). High school career academies and post-secondary outcomes. *Economics of Education Review*, 21(2), 137-152.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2), 698-714.

Meer, J. (2007). Evidence on the returns to secondary vocational education. *Economics of Education Review*, 26(5), 559-573.

Neumark, D., & Rothstein, D. (2006). School-to-career programs and transitions to employment and higher education. *Economics of Education Review*, 25(4), 374-393.

Page, L. C. (2012). Understanding the impact of career academy attendance: An application of the principal stratification framework for causal effects accounting for partial compliance. *Evaluation Review*, 36(2), 99-132.

Passarella, A. (2018). The Necessary Components of an Effective Career and Technical (CTE) Program.

Porter, J., & Yu, P. (2015). Regression discontinuity designs with unknown discontinuity points: Testing and estimation. *Journal of Econometrics*, 189(1), 132-147.

Table 1
Application Score Components

Year	Total Score	Language Arts	Math	Grades	Attendance	Leadership	Personal Statement
2006 Max (Weight)	100	20 (0.20)	20 (0.20)	40 (0.40)	20 (0.20)		
2007 Max (Weight)	100	21 (0.21)	21 (0.21)	48 (0.48)	10 (0.10)		
2008 Max (Weight)	100	21 (0.21)	21 (0.21)	48 (0.48)	10 (0.10)		
2009 Max (Weight)	106	21 (0.20)	21 (0.20)	48 (0.45)	10 (0.09)	3 (0.03)	3 (0.03)
2010 Max (Weight)	106	21 (0.20)	21 (0.20)	48 (0.45)	10 (0.09)	3 (0.03)	3 (0.03)
2011 Max (Weight)	120	21 (0.18)	21 (0.18)	48 (0.40)	10 (0.08)	10 (0.08)	10 (0.08)
2012 Max (Weight)	120	21 (0.18)	21 (0.18)	48 (0.40)	10 (0.08)	10 (0.08)	10 (0.08)
2013 Max (Weight)	120	21 (0.18)	21 (0.18)	48 (0.40)	10 (0.08)	10 (0.08)	10 (0.08)

Notes: Table presents overall admission score points and points associated with each component of the admission score for each application year in our sample. Numbers in parentheses represent the weight attached to each component when calculating the total application score.

Table 2
Summary Statistics

	State	CTHSS Full Sample		Around Threshold BW10	
	Mean	Mean	Std. Dev.	Mean	Std. Dev.
Female	0.49	0.42	0.49	0.42	0.49
Asian	0.04	0.01	0.11	0.01	0.10
Black	0.13	0.20	0.40	0.20	0.40
Hispanic	0.17	0.31	0.46	0.32	0.47
Free Lunch	0.33	0.65	0.48	0.67	0.47
English Learner	0.06	0.06	0.24	0.07	0.25
6^{th} Grade CMT-Reading	258	237	34.27	229	27.42
6^{th} Grade CMT-Math	262	238	34.89	230	27.21
6^{th} Grade CMT-Writing	255	230	33.43	224	27.72
Total Application Score	--	62.04	19.65	58.37	11.18
Application Grades Score	--	24.87	9.22	23.11	7.01
Application Attendance Score		6.95	4.27	6.96	4.17
Application Math Score	--	12.79	4.88	11.75	4.03
Application Language Arts Score	--	12.94	5.48	11.81	4.88

Notes: Table presents means and standard deviations of individual control variables, 6th grade standardized test scores, total application score and components of application score. Column 1 presents mean of characteristics for the state of Connecticut overall. Columns 2 and 3 present summary statistics for the full sample of students that applied to a CTHSS school. Columns 4 and 5 present summary statistics for the sample of students within +/- 10 points of the admission score threshold.

Table 3
Summary Statistics Above and Below Admission Threshold

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full Sample	Around Threshold BW10	Below Threshold BW10	Above Threshold BW10	Full Sample	Around Threshold BW10	Below Threshold BW10	Above Threshold BW10
Outcomes								
H.S. Graduation	0.83	0.83	0.77	0.87	0.83	0.82	0.78	0.85
2 Year College Enrollment	0.26	0.26	0.26	0.26	0.37	0.38	0.35	0.39
4 Year College Enrollment	0.23	0.18	0.16	0.20	0.40	0.32	0.27	0.35
Any College Enrollment	0.40	0.37	0.35	0.38	0.55	0.52	0.47	0.54
10th Grade Reading Score	218	212	205	217	224	216	208	220
10th Grade Math Score	238	231	221	236	225	215	205	221
10th Grade Writing Score	233	228	220	232	245	236	229	241
9th Grade Attendance	0.93	0.94	0.93	0.95	0.93	0.93	0.91	0.94
Total Earnings	63,117	63,890	55,669	68,762	43,984	44,790	42,972	45,900
Quarterly Earning	6,004	5,898	5,236	6,281	4,292	4,191	4,047	4,278
Quarters with Earnings	9.79	10.12	9.96	10.22	9.60	10.00	9.92	10.04
Control Variables								
Asian	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Black	0.17	0.17	0.19	0.16	0.23	0.25	0.26	0.24
Hispanic	0.27	0.27	0.29	0.27	0.36	0.40	0.40	0.39
Free Lunch	0.58	0.60	0.64	0.58	0.70	0.77	0.81	0.75
English Learner	0.05	0.06	0.07	0.05	0.07	0.09	0.10	0.08

Notes: Table presents summary statistics for main outcomes used in analysis and control variables. Columns 1-4 report means for sample of male students. Columns 5-8 report means for sample of female students. Columns 1 and 5 report means for full sample of students that applied to a CTHSS school. Columns 2-4 and 6-7 report means for students around the admission threshold based on a 10-point bandwidth and above and below the threshold.

Table 4
First-Stage Estimates BW10

	Probability of Being Admitted Full Sample		Probability of Attending Men		Probability of Attending Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome						
Offer	0.863***	0.863***	0.583***	0.583***	0.523***	0.524***
	(0.0208)	(0.0208)	(0.0228)	(0.0229)	(0.0238)	(0.0245)
Controls	No	Yes	No	Yes	No	Yes
F	1718	1727	656	649	485	458
Observations	15,339	15,339	9,287	9,287	7,629	7,629

Notes: Table presents first-stage estimates of the probability of being admitted to a CTHSS school and the probability of attending a CTHSS school. Columns 1 and 2 present first-stage estimates of the probability of being admitted to a CTHSS school where dependent variable is an indicator for receiving an offer of admittance and the sample includes both male and female students. Columns 3-4 present main first-stage estimates for probability of attending a CTHSS school after receiving an offer where dependent variable is an indicator for attendance at a CTHSS school in 9th grade and the sample is limited to male students. Columns 5 and 6 present the same information as columns 3 and 4 for the sample of female students. Specifications with controls include the full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5
Balancing Tests: Full Sample

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
	BW6	BW8	BW10	BW6	BW8	BW10
Centered Score	0.116***	0.0876***	0.0715***	0.114***	0.0878***	0.0713***
	(0.000631)	(0.000425)	(0.000319)	(0.000664)	(0.000465)	(0.000349)
Asian	-0.0286	-0.0187	-0.0122	-0.00570	-0.0140	0.00357
	(0.0235)	(0.0227)	(0.0221)	(0.0393)	(0.0325)	(0.0275)
Black	-0.00884	-0.0136	-0.00586	-0.00786	-0.00906	-0.00591
	(0.00975)	(0.00888)	(0.00826)	(0.00987)	(0.00902)	(0.00828)
Hispanic	-0.00765	-0.00963	-0.00569	-0.000398	0.00332	0.00123
	(0.00855)	(0.00785)	(0.00742)	(0.00910)	(0.00837)	(0.00771)
Free Lunch	0.00174	-2.22e-05	-0.000526	-0.00987	-0.0106	-0.00785
	(0.00748)	(0.00676)	(0.00631)	(0.00892)	(0.00829)	(0.00775)
English Learner	0.0160	0.0135	0.0146	-0.00694	-0.00829	-0.00283
	(0.0132)	(0.0119)	(0.0112)	(0.0121)	(0.0112)	(0.0104)
P-value	0.618	0.504	0.772	0.779	0.465	0.828
Observations	5,801	7,629	9,287	4,761	6,224	7,629

Notes: Table presents balancing tests for student characteristics. Columns 1-3 are based on the sample of male students that applied to a CTHSS school. Columns 4-6 are based on the sample of female students that applied to a CTHSS school. All specifications include CTHSS school-by-year fixed effects and resident fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6
Balancing Tests for Labor Market Sample

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
	BW6	BW8	BW10	BW6	BW8	BW10
Centered Score	0.115***	0.0875***	0.0715***	0.114***	0.0880***	0.0716***
	(0.000750)	(0.000511)	(0.000386)	(0.000821)	(0.000576)	(0.000435)
Asian	-0.0231	-0.00432	0.0151	-0.0292	-0.0491	-0.0200
	(0.0294)	(0.0281)	(0.0274)	(0.0494)	(0.0408)	(0.0339)
Black	-0.00463	-0.0158	-0.00706	-0.0128	-0.0136	-0.00961
	(0.0122)	(0.0109)	(0.0101)	(0.0125)	(0.0113)	(0.0105)
Hispanic	0.000247	-0.00361	-0.00127	-0.00608	-0.00391	-0.00444
	(0.0106)	(0.00976)	(0.00927)	(0.0119)	(0.0107)	(0.0100)
Free Lunch	0.00831	0.00431	0.00423	-0.00646	-0.0130	-0.00867
	(0.00889)	(0.00806)	(0.00751)	(0.0109)	(0.0101)	(0.00946)
English Learner	0.0104	0.00310	0.00945	-0.00650	-0.00317	-0.000879
	(0.0163)	(0.0149)	(0.0140)	(0.0154)	(0.0141)	(0.0131)
P-value	0.811	0.752	0.864	0.845	0.415	0.821
Observations	3,992	5,228	6,387	3,135	4,096	5,009

Notes: Table presents balancing tests for student characteristics based on the labor market sample. Columns 1-3 are based on the sample of male students that applied to a CTHSS school. Columns 4-6 are based on the sample of female students that applied to a CTHSS school. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 7
Reduced Form Estimates: Educational Outcomes BW10

Outcome	(1) Grad	(2) Grad	(3) Any Coll	(4) Any Coll	(5) 2Yr Coll	(6) 2Yr Coll	(7) 4Yr Coll	(8) 4Yr Coll
	Male Students							
Offer	0.0586***	0.0580***	-0.0459**	-0.0451**	-0.0263	-0.0261	-0.0336***	-0.0332***
	(0.0185)	(0.0183)	(0.0186)	(0.0194)	(0.0196)	(0.0197)	(0.0126)	(0.0126)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	9,284	9,284	9,287	9,287	7,788	7,788	9,287	9,287
	Female Students							
Offer	0.00656	0.00642	-0.0155	-0.0145	0.00259	0.00308	-0.0293	-0.0286
	(0.0208)	(0.0208)	(0.0257)	(0.0258)	(0.0310)	(0.0308)	(0.0265)	(0.0270)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	7,627	7,627	7,629	7,629	5,791	5,791	7,629	7,629

Notes: Table presents reduced form estimates for main nonlabor market outcomes. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Panel A presents estimates based on sample of male students. Panel B presents the same information for female students. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8
Reduced Form Estimates: Labor Market Outcomes BW10

Outcome	(1) Total Earnings	(2) Total Earnings	(3) Quarterly Earnings	(4) Quarterly Earnings	(5) Quarters with Earnings	(6) Quarters with Earnings	(7) Quarterly Earnings at 23	(8) Quarterly Earnings at 23
	Male Students							
Offer	0.213***	0.212***	0.187***	0.185***	0.589**	0.596**	0.204***	0.201***
	(0.0473)	(0.0479)	(0.0357)	(0.0347)	(0.257)	(0.261)	(0.0599)	(0.0586)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	6,387	6,387	4,980	4,980	6,387	6,387	1,929	1,929
	Female Students							
Offer	0.0110	0.0101	0.0185	0.0173	0.00482	0.00395	0.0343	0.0213
	(0.0602)	(0.0615)	(0.0337)	(0.0346)	(0.249)	(0.246)	(0.0486)	(0.0510)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,009	5,009	3,919	3,919	5,009	5,009	1,441	1,441

Notes: Table presents reduced form estimates for labor market outcomes. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Panel A presents reduced form estimates based on sample of male students in labor market. Panel B presents the same information for the sample of female students with labor market outcomes. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9
2SLS Estimates Male Students BW10

Outcome	(1) Grad	(2) Grad	(3) Any Coll	(4) Any Coll	(5) 2Yr Coll	(6) 2Yr Coll	(7) 4Yr Coll	(8) 4Yr Coll
Attend	0.100***	0.0996***	-0.0786**	-0.0774**	-0.0442	-0.0439	-0.0576***	-0.0569***
	(0.0332)	(0.0328)	(0.0317)	(0.0332)	(0.0328)	(0.0332)	(0.0214)	(0.0215)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	9,284	9,284	9,287	9,287	7,788	7,788	9,287	9,287
Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings at 23	Quarterly Earnings at 23
Attend	0.366***	0.364***	0.313***	0.309***	1.011**	1.024**	0.316***	0.312***
	(0.0792)	(0.0792)	(0.0616)	(0.0594)	(0.431)	(0.438)	(0.0952)	(0.0930)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	6,387	6,387	4,980	4,980	6,387	6,387	1,929	1,929

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10
Mechanisms: Reduced Form and 2SLS Male Students BW10

	Reduced Form Estimates		2SLS Estimates	
	(1)	(2)	(3)	(4)
Outcome	9th Grade Attendance	10th Grade Test Scores	9th Grade Attendance	10th Grade Test Scores
Offer	0.0100***	0.109***	0.0172***	0.177***
	(0.00307)	(0.0300)	(0.00529)	(0.0474)
Controls	Yes	Yes	Yes	Yes
Observations	9,287	6,257	9,287	6,257

Notes: Table presents reduced-form and 2SLS RD estimates for 9th grade days of attendance and standardized composite 10th grade reading and math test scores. Sample is limited to male students. All estimates based on local linear regression and a 10-point bandwidth. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11
Alternative Bandwidth Male Students

Outcome	(1) Grad	(2) Any Coll	(3) Total Earnings	(4) Quarterly Earnings	(5) Quarters with Earnings	6) Quarterly Earnings at 23
			Bandwidth 6			
Attend	0.0588**	-0.0789***	0.272***	0.210***	1.047**	0.239***
	(0.0256)	(0.0238)	(0.0758)	(0.0536)	(0.448)	(0.0651)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,800	5,801	3,992	3,101	3,992	1,224
			Bandwidth 8			
Attend	0.0553***	-0.0592***	0.296***	0.209***	0.847**	0.220***
	(0.0199)	(0.0174)	(0.0635)	(0.0377)	(0.377)	(0.0644)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,626	7,629	5,228	4,083	5,228	1,588
			Bandwidth 10			
Attend	0.0580***	-0.0451**	0.212***	0.185***	0.596**	0.201***
	(0.0183)	(0.0194)	(0.0479)	(0.0347)	(0.261)	(0.0586)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,284	9,287	6,387	4,980	6,387	1,929

Notes: Table presents reduced-form RD estimates for main outcomes based on various bandwidths. Sample is limited to male students. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12
Reduced Form Falsification Tests Male Students BW10

Outcome	(1) Grad	(2) Any College	(3) Total Earnings	(4) Quarterly Earnings	(5) Quarters with Earnings	(6) Quarterly Earnings at 23
			Cutoff -10 Points			
Offer	0.0200	0.00407	0.0153	-0.000233	0.0826	-0.0554
	(0.0236)	(0.0239)	(0.0656)	(0.0433)	(0.245)	(0.0598)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,235	6,236	4,008	3,070	4,008	1,109
			Cutoff + 10 Points			
Offer	0.0132	-0.0277	0.0764	0.0413	0.199	-0.00500
	(0.0116)	(0.0182)	(0.0500)	(0.0323)	(0.178)	(0.0533)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,391	11,395	7,820	6,205	7,820	2,451
			Cutoff + 15 Points			
Offer	-0.00274	-0.00949	0.00295	-0.0309	-0.0405	0.0318
	(0.00960)	(0.0189)	(0.0592)	(0.0324)	(0.193)	(0.0478)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,050	10,053	6,893	5,459	6,893	2,159
			Cutoff + 20 Points			
Offer	0.00346	0.0231	-0.0393	-0.00447	-0.172	-0.0623*
	(0.00819)	(0.0198)	(0.0599)	(0.0234)	(0.229)	(0.0350)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,878	8,880	6,066	4,804	6,066	1,856

Notes: Table presents reduced-form RD falsification tests for main outcomes based on pseudo cutoffs where we move the actual cutoff threshold: 1) down 10 points, 2) up 10 points, 3) up 15 points, and 4) up 20 points. Sample is limited to male students. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 13
Heterogeneity Analysis Main Outcomes Male Students BW10

Outcome	Graduation	Any College	Quarterly Earnings	Quarters with Earnings
	(1)	(2)	(3)	(4)
A. Free Lunch Status				
Attend	0.0900**	-0.0497	0.328***	1.608**
	(0.0409)	(0.0582)	(0.0947)	(0.768)
Attend*Free Lunch	0.0131	-0.0422	-0.0327	-0.833
	(0.0473)	(0.0771)	(0.113)	(0.909)
Controls	Yes	Yes	Yes	Yes
Observations	9,284	9,287	4,980	6,387
B. Race/Ethnicity				
Attend	0.124***	-0.0936*	0.380***	0.512
	(0.0415)	(0.0530)	(0.0973)	(0.593)
Attend*Black/Hispanic	-0.0439	0.0264	-0.172	0.843
	(0.0584)	(0.0770)	(0.113)	(0.813)
Controls	Yes	Yes	Yes	Yes
Observations	9,284	9,287	4,980	6,387
C. Central City				
Attend	0.104**	-0.0798*	0.342***	0.506
	(0.0449)	(0.0463)	(0.0761)	(0.431)
Attend*Central City	-0.0187	0.00538	-0.137	1.518**
	(0.0704)	(0.0549)	(0.130)	(0.580)
Controls	Yes	Yes	Yes	Yes
Observations	9,284	9,287	4,980	6,387

Notes: Table presents 2SLS RD estimates for main outcomes based on sample of male students. Panel A interacts the attend a CTE school indicator with the indicator for free lunch eligible students. Panel B interacts the attend indicator with an indicator for whether a student is Black or Hispanic. Panel C interacts the attend indicator with an indicator for whether a student resides in one of Connecticut's central cities. All estimates based on local linear regression and a 10-point bandwidth. All specifications include the full set of controls listed in Table 3, CTHSS school-by-year fixed effects and resident town fixed effects and interactions between the relevant share of elective courses and the running variable and the running variable interacted with offer. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 14
CTE Course Summary Statistics

	CTHSS Schools		Non-CTHSS Schools	
	Mean	Std. Dev.	Mean	Std. Dev.
Share Total CTE Courses	0.891	0.081	0.453	0.067
Share Trade CTE Courses	0.567	0.113	0.054	0.023
Share Human Services/Tourism Hospitality	0.067	0.027	0.088	0.027

Notes: Table presents means and standard deviations of the share of elective courses that are CTE courses. Columns 1 and 2 presents summary statistics for CTHSS schools while columns 3 and 4 present summary statistics for non-CTHSS schools. The first row shows the share of elective courses that are any type of CTE course. The second row shows the share of elective courses that are CTE trade courses (architecture, transportation and manufacturing). The third row shows the share of elective courses that are human service, tourism and hospitality or Family & Consumer Sciences. See appendix for definitions of various CTE courses.

Table 15
Counterfactual CTE Courses 2SLS BW10 Male Students

Outcome	Graduation	Any College	Quarterly Earnings	Quarters with Earnings
	(1)	(2)	(3)	(4)
A. All CTE Courses				
Offer	0.108***	-0.0706**	0.331***	1.132**
	(0.0355)	(0.0311)	(0.0628)	(0.477)
Attend*CTE	-0.00958	0.0685*	-0.169**	0.0884
	(0.0386)	(0.0358)	(0.0721)	(0.478)
Controls	Yes	Yes	Yes	Yes
Observations	9,180	9,183	4,911	6,311
B. Trade Courses				
Attend	0.0994***	-0.0700**	0.329***	1.192***
	(0.0340)	(0.0323)	(0.0561)	(0.445)
Attend*Trade	-0.0717*	0.0590*	-0.193***	-0.0761
	(0.0363)	(0.0316)	(0.0598)	(0.566)
Controls	Yes	Yes	Yes	Yes
Observations	9,180	9,183	4,911	6,311
C. Human Services/Tourism Hospitality				
Attend	0.106***	-0.0796**	0.315***	1.148**
	(0.0343)	(0.0324)	(0.0637)	(0.472)
Attend*HS/TH	0.00210	-0.0114	-0.0142	0.0826
	(0.0309)	(0.0276)	(0.0509)	(0.428)
Controls	Yes	Yes	Yes	Yes
Observations	9,180	9,183	4,911	6,311

Notes: Table presents 2SLS RD estimates for main outcomes based on sample of male students. Panel A interacts the attend a CTE school indicator with the share of elective courses that are CTE in a student's resident school district. Panel B interacts the attend indicator with the share of elective courses that are trade courses (architecture, transportation and manufacturing) in a student's resident school district while Panel C interacts the attend indicator with the share of elective courses that are human service, tourism and hospitality or Family & Consumer Sciences. All specifications include the full set of controls listed in Table 3, CTHSS school-by-year fixed effects and resident town fixed effects and interactions between the relevant share of elective courses and the running variable and the running variable interacted with offer. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 1A: Probability of Being Admitted to a CTHSS School Full Sample

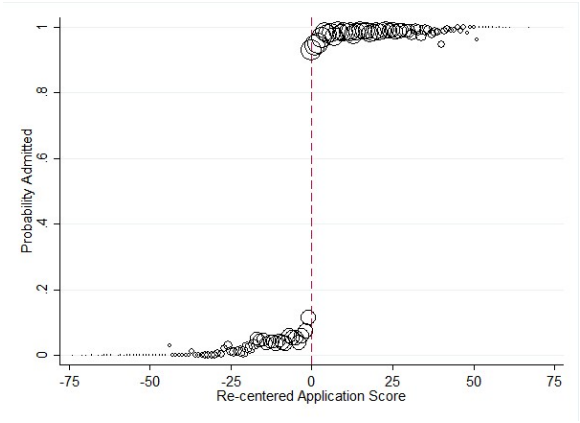


Figure 1B: Probability of Attending a CTHSS School Men

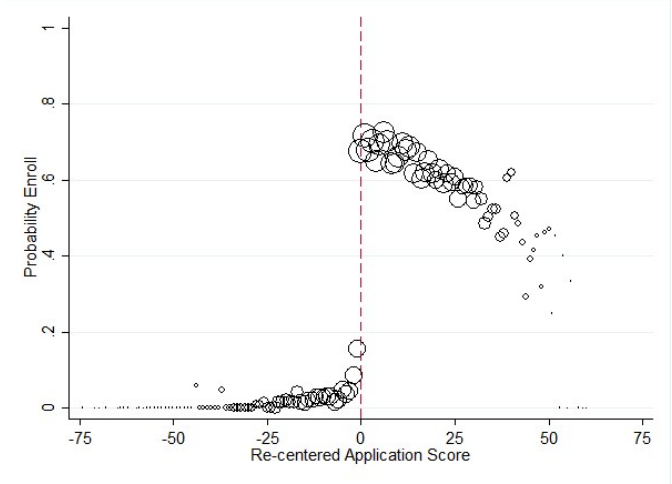


Figure 1C: Probability of Attending a CTHSS School Women

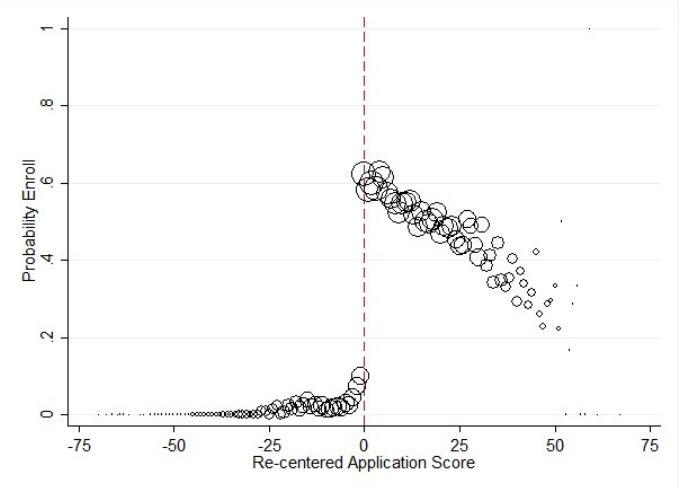
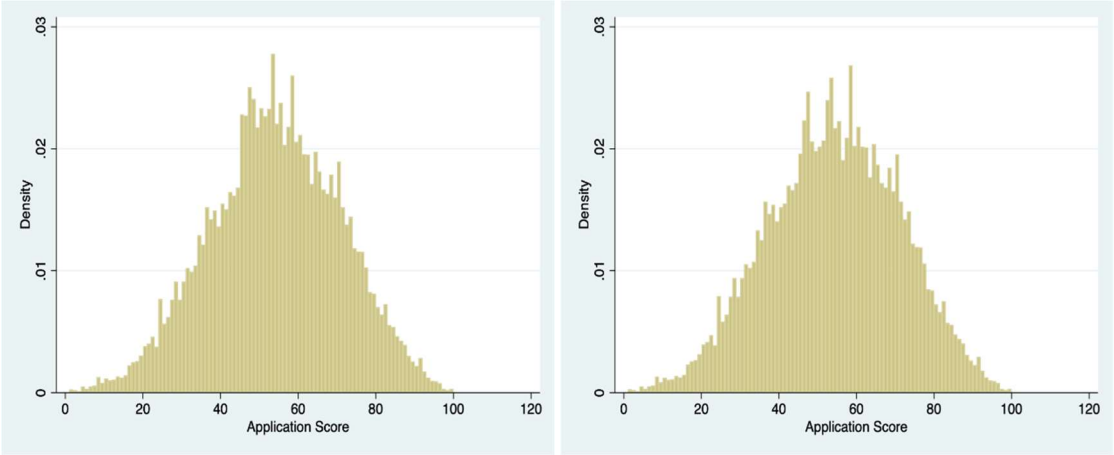


Figure 2: Uconditional Distribution of Running Variable

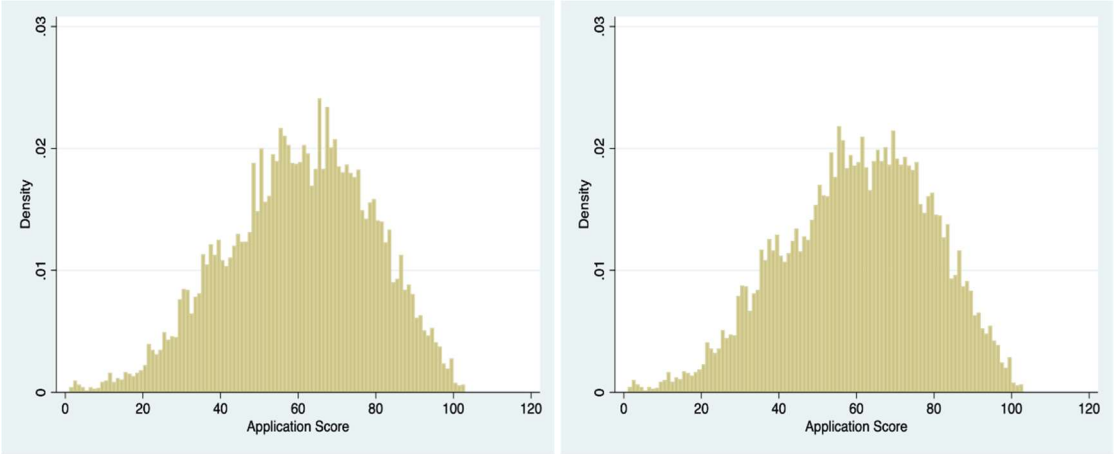
2006-2008

Actual Distribution Dropping Observations at the Cut-off



2009-2010

Actual Distribution Dropping Observations at the Cut-off



2011-2013

Actual Distribution Dropping Observations at the Cut-off

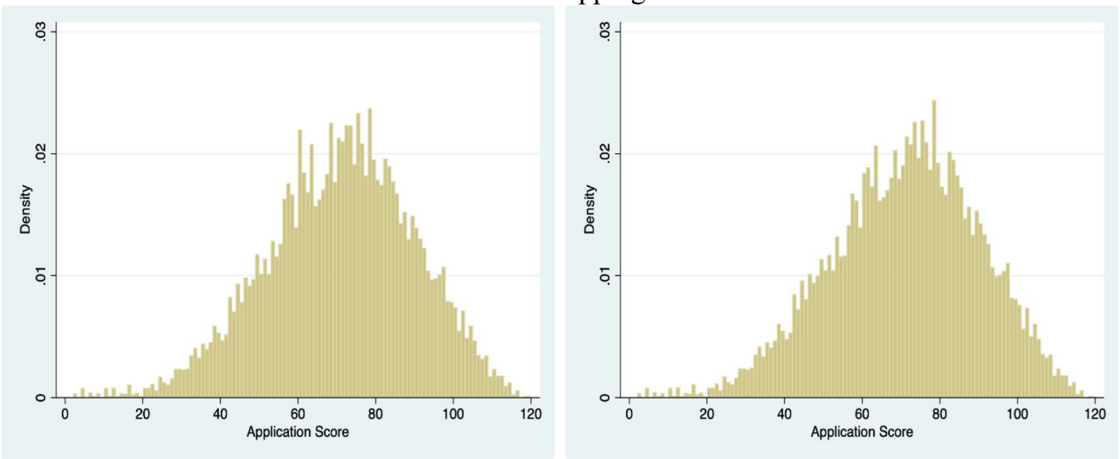
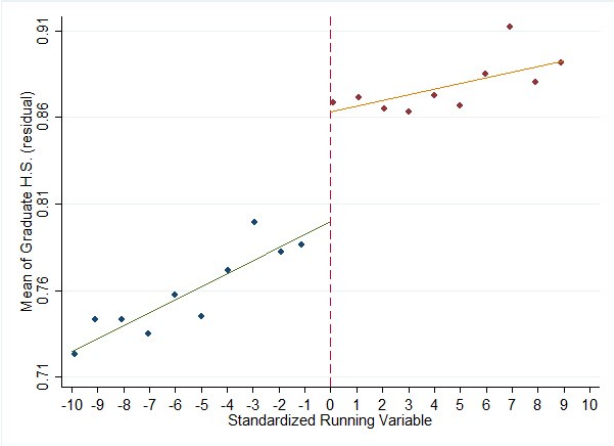
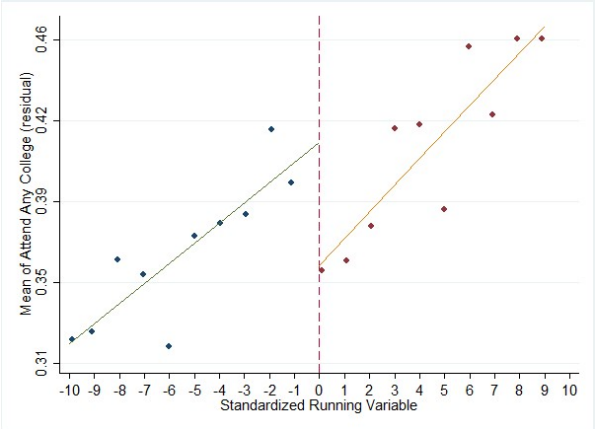


Figure 3: Reduced Form Graphs H.S. Graduation and Attend College

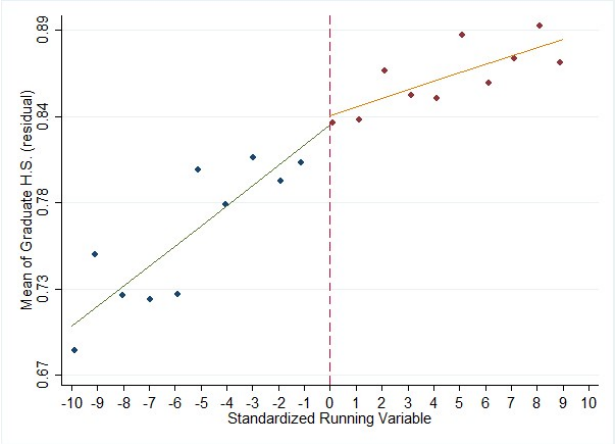
High School Graduation Men



Attend College Men



High School Graduation Women



Attend College Women

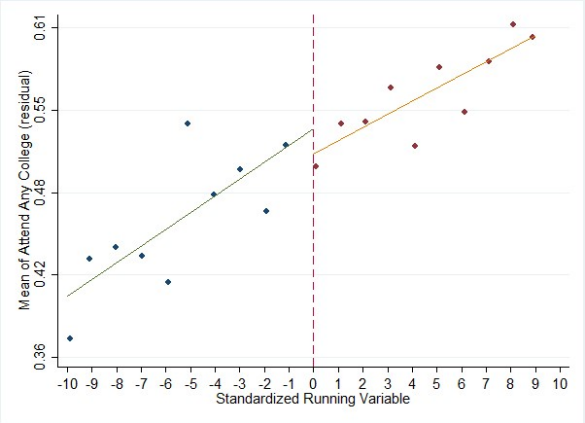
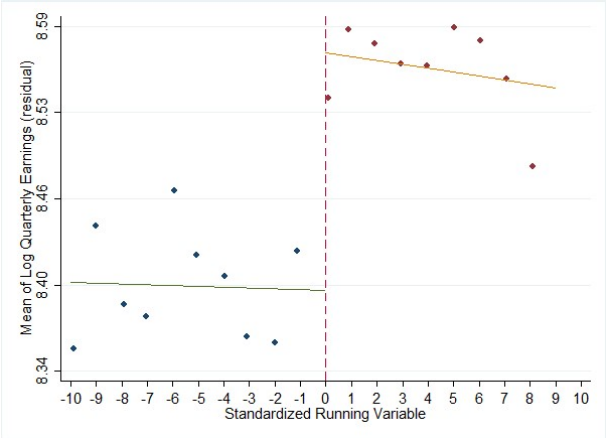
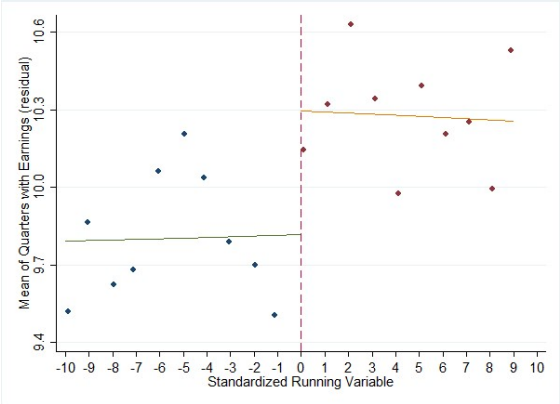


Figure 4: Reduced Form Quarterly Earnings and Quarters with Earnings

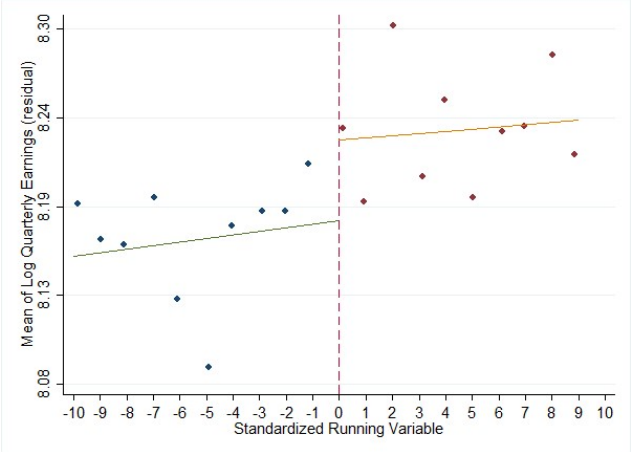
Log Quarterly Earnings Men



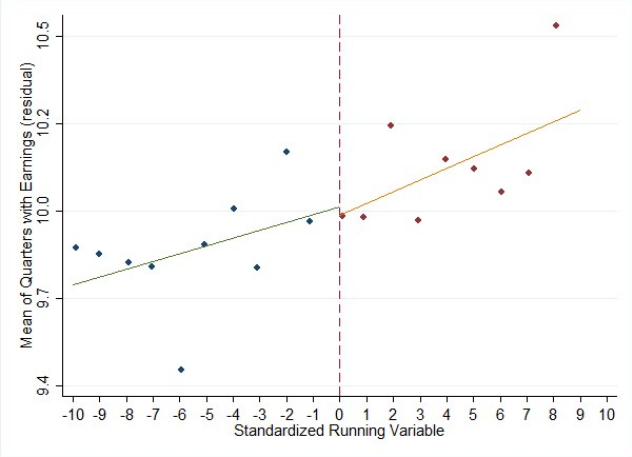
Quarters with Earnings Men



Log Quarterly Earnings Women



Quarters with Earnings Women



Empirical Appendix Tables

Table 1A
First-Stage Estimates Labor Market Sample BW10

	Probability of Attending Men		Probability of Attending Women	
	(1)	(2)	(3)	(4)
Outcome				
Offer	0.582***	0.582***	0.538***	0.539***
	(0.0266)	(0.0269)	(0.0289)	(0.0296)
Controls	No	Yes	No	Yes
F	480	467	345	332
Observations	6,387	6,387	5,009	5,009

Notes: Table presents first-stage estimates based on sample of students in labor market. Estimates are for probability of attending a CTHSS school after receiving an offer where dependent variable is an indicator for attendance at a CTHSS school in 9th grade. Columns 1 and 2 are for sample of male students in labor market. Columns 3 and 4 are for sample of female students in labor market. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2A
Robustness of Core Results to Dropping Application Years 2010 and Later

	(1)	(2)	(3)	(4)	(5)
Outcome	Grad	Any Coll	Quarterly Earnings	Quarters With Earnings	Quarterly Earnings at 23
	Main Reduced Form Results from Paper Sample of Male Students				
Attend	0.0580***	-0.0451**	0.185***	0.596**	0.201***
	(0.0183)	(0.0194)	(0.0347)	(0.261)	(0.0586)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	9,284	9,287	4,980	6,387	1,929
	Reduced Form Results Based on 2006-2009 Sample of Male Students				
Attend	0.0618**	-0.00812	0.147***	1.038**	0.200***
	(0.0249)	(0.0292)	(0.0347)	(0.424)	(0.0586)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	4,515	4,515	3,539	3,846	1,924

Notes: Table presents reduced-form RD estimates for main outcomes based on the sample of male students. Panel A replicates results presented in main paper for comparison purposes. Panel B limits the sample to application years 2006 - 2009. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3A
Balancing Tests: No Donut Hole

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
	BW6	BW8	BW10	BW6	BW8	BW10
Centered Score	0.112***	0.0859***	0.0705***	0.111***	0.0860***	0.0700***
	(0.000703)	(0.000458)	(0.000338)	(0.000740)	(0.000501)	(0.000371)
Asian	-0.0364	-0.0233	-0.0162	-0.0172	-0.0277	-0.00952
	(0.0267)	(0.0250)	(0.0236)	(0.0432)	(0.0349)	(0.0291)
Black	-0.00405	-0.00959	-0.00287	-0.00482	-0.00598	-0.00357
	(0.0104)	(0.00935)	(0.00862)	(0.0107)	(0.00960)	(0.00871)
Hispanic	-0.00980	-0.0103	-0.00625	0.000454	0.00523	0.00294
	(0.00914)	(0.00820)	(0.00768)	(0.00988)	(0.00889)	(0.00807)
Free Lunch	-0.00619	-0.00587	-0.00487	0.00345	-0.000666	0.000235
	(0.00809)	(0.00714)	(0.00659)	(0.00965)	(0.00875)	(0.00808)
English Learner	0.0195	0.0157	0.0161	0.000690	-0.00216	0.00125
	(0.0141)	(0.0125)	(0.0116)	(0.0133)	(0.0120)	(0.0110)
P-value	0.425	0.450	0.650	0.986	0.781	0.967
Observations	6,553	8,381	10,039	5,370	6,832	8,236

Notes: Table presents balancing tests for student characteristics based on specifications without a donut hole. Columns 1-3 are based on the sample of male students that applied to a CTHSS school. Columns 4-6 are based on the sample of female students that applied to a CTHSS school. All specifications include CTHSS school-by-year fixed effects and resident fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4A
Individual Covariate Balancing Tests

Outcome	(1) Asian	(2) Black	(3) Hispanic	(4) Free Lunch	(5) English Learner
			Bandwidth 6		
Offer	-0.00354	-0.0135	-0.00774	0.00435	0.0175
	(0.00635)	(0.0166)	(0.0189)	(0.0229)	(0.0136)
Observations	5,801	5,801	5,801	5,801	5,801
			Bandwidth 8		
Offer	-0.00298	-0.0186	-0.00654	-0.000910	0.0173
	(0.00475)	(0.0132)	(0.0106)	(0.0150)	(0.0157)
Observations	7,629	7,629	7,629	7,629	7,629
			Bandwidth 10		
Offer	-0.00215	-0.00503	-0.00612	-0.00283	0.0158
	(0.00411)	(0.0126)	(0.0106)	(0.0152)	(0.0124)
Observations	9,287	9,287	9,287	9,287	9,287

Notes: Table presents balancing tests for individual student characteristics based on sample of male students. Estimates are from a RD specification using local linear regression and the bandwidth listed at the top of each panel. All specifications include CTHSS school-by-year fixed effects and resident fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 5A
Individual Covariate Balancing Tests No Donut Hole Men

Outcome	(1) Asian	(2) Black	(3) Hispanic	(4) Free Lunch	(5) English Learner
			Bandwidth 6		
Offer	-0.00156	-0.00185	-0.0185	-0.00883	0.0148
	(0.00622)	(0.0176)	(0.0152)	(0.0221)	(0.0116)
Observations	6,553	6,553	6,553	6,553	6,553
			Bandwidth 8		
Offer	-0.00340	-0.00711	-0.0119	-0.00920	0.0177
	(0.00450)	(0.0128)	(0.00832)	(0.0152)	(0.0123)
Observations	8,381	8,381	8,381	8,381	8,381
			Bandwidth 10		
Offer	-0.00276	0.00231	-0.0111	-0.0103	0.0158
	(0.00379)	(0.0124)	(0.00790)	(0.0156)	(0.00986)
Observations	10,039	10,039	10,039	10,039	10,039

Notes: Table presents balancing tests for individual student characteristics based on sample of male students and no donut hole. Estimates are from a RD specification using local linear regression and the bandwidth listed at the top of each panel. All specifications include CTHSS school-by-year fixed effects and resident fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6A
Reduced Form Estimates Grad and College: No Donut Hole BW10

Outcome	(1) Grad	(2) Grad	(3) Any Coll	(4) Any Coll	(5) 2Yr Coll	(6) 2Yr Coll	(7) 4Yr Coll	(8) 4Yr Coll
	Male Students							
Offer	0.0621***	0.0611***	0.0505***	0.0502***	-0.0347*	-0.0351*	-0.0297**	-0.0298**
	(0.0179)	(0.0177)	(0.0179)	(0.0181)	(0.0182)	(0.0180)	(0.0119)	(0.0120)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	10,036	10,036	10,039	10,039	8,425	8,425	10,039	10,039
	Female Students							
Offer	0.00411	0.00446	-0.0197	-0.0188	-0.00580	-0.00585	-0.0257	-0.0252
	(0.0177)	(0.0180)	(0.0236)	(0.0235)	(0.0287)	(0.0286)	(0.0255)	(0.0258)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	8,234	8,234	8,236	8,236	6,264	6,264	8,236	8,236

Notes: Table presents reduced form estimates for main nonlabor market outcomes based on specifications without a donut hole. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Panel A presents estimates based on sample of male students. Panel B presents the same information for female students. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7A
Reduced Form Estimates for Educational Outcomes Based on Labor Market Sample Men

Outcome	(1) Grad	(2) Grad	(3) Any Coll	(4) Any Coll	(5) 2Yr Coll	(6) 2Yr Coll	(7) 4Yr Coll	(8) 4Yr Coll
	Donut Hole Sample of Male Students							
Offer	0.0560***	0.0564***	-0.0524*	-0.0528*	-0.0406	-0.0417	-0.0250*	-0.0247*
	(0.0172)	(0.0165)	(0.0267)	(0.0269)	(0.0290)	(0.0288)	(0.0146)	(0.0146)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	6,387	6,387	6,387	6,387	5,334	5,334	6,387	6,387

Notes: Table presents reduced form estimates for main non-labor market outcomes based on sample of male students in labor market. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8A
Reduced Form Estimates Labor Market Outcomes: No Donut Hole BW10

Outcome	(1) Total Earnings	(2) Total Earnings	(3) Quarterly Earnings	(4) Quarterly Earnings	(5) Quarters with Earnings	(6) Quarters with Earnings	(7) Quarterly Earnings at 23	(8) Quarterly Earnings at 23
	Male Students							
Offer	0.209***	0.213***	0.171***	0.170***	0.476**	0.487**	0.160**	0.161***
	(0.0479)	(0.0484)	(0.0350)	(0.0335)	(0.233)	(0.237)	(0.0612)	(0.0595)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	6,913	6,913	5,426	5,426	6,913	6,913	2,081	2,081
	Female Students							
Offer	0.0111	0.0119	0.0302	0.0312	-0.0182	-0.0194	0.0321	0.0223
	(0.0657)	(0.0670)	(0.0319)	(0.0323)	(0.249)	(0.246)	(0.0421)	(0.0438)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	5,394	5,394	4,228	4,228	5,394	5,394	1,552	1,552

Notes: Table presents reduced form estimates for labor market outcomes based on specifications without a donut hole. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Panel A presents reduced form estimates based on sample of male students in labor market. Panel B presents the same information for the sample of female students with labor market outcomes. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9A
2SLS Estimates Male Students: No Donut Hole BW 10

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Grad	Grad	Any Coll	Any Coll	2Yr Coll	2Yr Coll	4Yr Coll	4Yr Coll
Attend	0.109***	0.108***	-0.0889***	-0.0885***	-0.0601*	-0.0608*	-0.0522**	-0.0525**
	(0.0328)	(0.0325)	(0.0311)	(0.0316)	(0.0312)	(0.0310)	(0.0206)	(0.0209)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	10,036	10,036	10,039	10,039	8,425	8,425	10,039	10,039
Outcome	Total Earnings	Total Earnings	Quarterly Earnings	Quarterly Earnings	Quarters with Earnings	Quarters with Earnings	Quarterly Earnings at 23	Quarterly Earnings at 23
Attend	0.365***	0.372***	0.290***	0.288***	0.834**	0.853**	0.268**	0.270**
	(0.0794)	(0.0800)	(0.0584)	(0.0559)	(0.398)	(0.406)	(0.105)	(0.103)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	6,913	6,913	5,426	5,426	6,913	6,913	2,081	2,081

Notes: Table presents 2SLS estimates for main outcomes based on sample of male students and no donut hole. All estimates are based on a RD specification using local linear regression and a 10-point bandwidth. Columns with controls include full set of controls listed in Table 3. All specifications include CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10A
Fisher Exact Tests

Outcome	Graduation	Any College	Quarterly Earnings	Quarters with Earnings
	(1)	(2)	(3)	(4)
		p-values for Core Results		
Offer	0.002	0.015	0.000	0.024
		Fisher Exact Test *p*-values		
Offer	0.000	0.000	0.000	0.000
Observations	9,284	9,287	4,980	6,387

Notes: Table presents reduced-form RD *p*-values for estimated coefficient on offer. Top panel presents *p*-values associated with reduced form estimates presented in Tables 7 and 8. Bottom panel presents *p*-values associated with Fisher's exact test as proposed by Cattaneo et al. (2019).

Table 11A
Falsification Test: First Stage Male Students BW10

Outcome	(1) Cutoff -10	(2) Cutoff +10	(3) Cutoff +15	(4) Cutoff +20
A. Full Sample				
Offer	-0.00508	0.0173	0.00524	0.0220
	(0.00682)	(0.0194)	(0.0228)	(0.0186)
Controls	Yes	Yes	Yes	Yes
Observations	5,885	11,395	10,053	8,880
B. Labor Market Sample				
Offer	-0.00206	0.0157	0.00862	0.0370
	(0.00722)	(0.0238)	(0.0266)	(0.0253)
Controls	Yes	Yes	Yes	Yes
Observations	4,008	7,820	6,893	6,066

Notes: Table presents first-stage RD falsification tests based on pseudo cutoffs where we move the actual cutoff threshold: 1) down 10 points, 2) up 10 points, 3) up 15 points, and 4) up 20 points. Panel A is for the full set of male students that applied to a CTHSS school. Panel B is for sample of male students in labor market. All specifications include the full set of controls listed in Table 3 and CTHSS school-by-year fixed effects and resident town fixed effects. Robust standard errors, clustered at the school-by-year and town levels in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Methodological Appendix

As discussed in the paper, the Connecticut Technical High School System (CTHSS) admits students based on a discrete score variable that runs between zero and a maximum of between 100 and 120 depending upon the year considered. This score is the sum of individual components based on standardized state level mathematics and language arts test scores, grade point average, attendance, and for 2009 or later, points assigned based on extracurricular activities and a written statement. The weights assigned to each component vary from year to year, and the exact weights are shown in Table 1 in the paper itself.

Two issues arise, however, that complicate our analysis relative to a simple, fuzzy regression discontinuity. First, the individual schools never establish a formal cut-off score, and in fact we can find many examples of individuals receiving acceptance letters who have scores below the scores of individuals who did not receive acceptance letters for the same school and application year. Therefore, there are two sources of noise associated with the discontinuity: deviations of school administrators from a high to low score admission/acceptance rule, and differences in take-up among accepted applicants, and the true location of the discontinuity is unknown. The second issue that arises is that the individual score components are both correlated with each other and discrete in nature leading to raw score distributions that have heaps or mass points at specific locations in the distribution, making it impossible to directly identify manipulation at the threshold or boundary using standard McCrary tests.

1. Empirical Identification of Cut-offs

While receipt of an acceptance letter is not perfectly determined by a student's application score, the mass of the acceptances for each school and year appear concentrated above apparent score thresholds, and this concentration is especially true for the initial batch of acceptance letters sent out by the schools each year. Therefore, we follow Porter and Yu (2015) and select the threshold for each year and school from the empirical distribution of applications.

Porter and Yu (2015) recommend estimating the location of the threshold by selecting the threshold that maximizes the size of the discontinuity, or specifically:

$$\widehat{X^*} = argmax_{X^*}\hat{\alpha}(X^*) \tag{A1}$$

where the treatment (T) is defined by

$$T = \begin{cases} 1 & if\ y^* \geq 0 \\ 0 & if\ y^* < 0 \end{cases},\ T^* = \begin{cases} f_1(X) + \varepsilon & if\ X < X^* \\ f_2(X) + \alpha + \varepsilon & if\ X \geq X^* \end{cases},\ f_1(X^*) = f_2(X^*) \tag{A2}$$

and $f_1(X)$ and $f_2(X)$ are continuous and differential functions. The inclusion of different functions on either side of the threshold allows for differential processes for non-compliance. In practice, we estimate equation (A2) using a linear probability model and specifying $f_1(X)$ and $f_2(X)$ as linear functions of X.

Porter and Yu (2015) estimate $\widehat{X^*}$ using the discontinuity in the outcome in order to develop second stage specification tests for the existence of treatment effects at that unknown discontinuity. However, Porter and Yu's specification tests require both continuous outcomes and a smooth distribution of the population over scores X. Therefore, we take a different approach exploiting fuzzy regression discontinuity approaches. We select the thresholds applying equation (A1) to the actual school decision to send an acceptance letter to an applicant since that is the decision process that creates the discontinuity. Then, with the estimated thresholds in hand, we estimate a two stage least squares model of student outcomes (y) that incorporates the additional noise created by student take-up so that the first stage is whether we see the applicant in the technical high school the year after they applied (A). Specifically, the fuzzy RD takes the following form:

$$y = \beta A + g_1(X) + d(\widehat{X^*} \leq X)g_2(X) + \varepsilon \tag{A3}$$

where A is instrumented by

$$A = \tilde{\alpha} d(\widehat{X^*} \leq X) + h_1(X) + d(\widehat{X^*} > X)h_2(X) + \varepsilon \tag{A4}$$

$d(\widehat{X^*} \leq X)$ is an indicator function that takes the value one when the condition is satisfied, and both $g_1(\widehat{X^*}) = g_2(\widehat{X^*})$ and $h_1(\widehat{X^*}) = h_2(\widehat{X^*})$.

The disadvantage of this approach is that our analysis is conditional on the estimated discontinuity and the estimated discontinuity may not even represent a true discontinuity. However, in a 2SLS context, we only need to establish the power of the instrument and the validity of the exclusion restriction. The issue of whether a discontinuity exists can be determined by examining the power of the instrument. If the estimated threshold has significant power to explain student attendance A then the discontinuity exists, and the second stage estimates and standard errors from 2SLS will be consistent as long as the exclusion restriction is valid. Determining the power of the instrument, however, is not straightforward since the cut-off has been selected using the same data that is then used to estimate the first stage equation (A4). Naturally, treatment (receiving an acceptance letter) affects attendance and so the *F*-test may be biased upwards providing misleading evidence on instrument power because we

selected $\widehat{X^*}$ to maximize the discontinuity for T which is strongly correlated with A. We follow Card, Mas and Rothstein (2008) and address this concern using a hold-out sample.[1]

1.1 Selection of Cut-offs using Hold-out Samples

In our case, we will divide the applicants in each school and year into equal sized analysis and hold-out samples so that we can use the analysis sample to select the thresholds and use the hold-out sample to estimate the first stage for attendance and examine the power of the instrument. Specifically, we do the following:

1) For each school and year, divide all applicants into deciles, assign each applicant a random number, place applicants in the hold-out sample if they are above the median on the random number and in the analysis sample if below, split median applicants half in hold-out half in analysis samples, and in the case of an odd number of median applicants assign the last median applicant randomly to either the hold-out or analysis sample.
2) Using our preferred bandwidth of 10 points and linear running variables, estimate equation A1 for each school and year starting with a candidate cut-off score at 10 so that the bottom of the 20-point band is at a score of zero using the analysis sample.
3) Re-estimate these models incrementing the candidate cut-off score by 1 each time and ending 10 points away from the maximum score so that the top of the band at the maximum.
4) Select the cut-off for each school and year as the cut-off that provides the maximum estimate of α.[2]
5) Center the scores for each school and year by subtracting the cut-off and pool all years and schools.
6) Estimate equation (A4) for school attendance and calculate the *F*-statistic associated with the indicator $d(\widehat{X^*} \leq X)$ using the hold-out sample.

We conduct this hold-out simulation four times, and the resulting *F*-statistics always fall between 456 and 628 for the full sample and between 458 and 674 for the donut hole sample. We also estimate the first stage models separately for each school and year. The means of the estimated discontinuities over all schools and years range between 0.525 and 0.540 for the four simulations very close to our full sample

[1] Card, Mas and Rothstein (2008) calculate tipping points associated with racial segregation for each metropolitan area using a subset of census tracts, and then analyzed the effects of being above the tipping point using the remaining tracts.
[2] Note that we verify that the cut-offs are never anywhere near the edge of the score range considered.

first stage estimate of 0.582. The fraction of thresholds that are significant at the 10 percent level ranges between 0.795 and 0.843. Further, the average threshold size across years for any school or simulation never falls below 0.329, and the fraction of significant thresholds for any school in a simulation never falls below 0.571. However, while the empirically selected thresholds provide a very strong instrument for explaining school attendance across the sample of CTHSS schools, the magnitude of the estimated threshold varies considerably and systematically across the simulations. The correlation between the average threshold for each school across the simulations is very high ranging between 0.77 and 0.91. We observe less evidence of a systematic difference in thresholds across years with the correlation between the average threshold for each year ranging between -0.10 and 0.56. Appendix Table 1B presents the correlations of the average thresholds by school across the four simulations in panel 1, and by year in panel 2. Appendix Table 2B presents the full set of correlations for one of the simulations as an example.

2. Examining the Empirical Distribution of Scores

While the underlying test scores, grade point averages and attendance data are relatively continuous, the system for assigning points divides students into more aggregated bins leading to mass points in the empirical distribution of application scores. The raw distribution of applicant admission scores are shown in the left hand side of Figure 2 of the paper separately for 2006 to 2008, 2009-2010 and 2011 to 2014 since prior to 2008 the score did not include points from the written statement and extracurricular activities, and in 2011 those points increased from a total of 6 points to a total of 120 points. The distribution contains significant mass points as well as holes in the distribution, which would typically raise concerns about manipulation at the cut-off. However, the right-hand side of Figure 2 shows the same distributions after dropping students for each school and application year whose score was exactly at the cut-off and dropping those students has at most modest effects on the smoothness of the raw score distribution. Similarly, the top panel of Appendix Figures 1B-8B present the same raw score distributions separately for each application year with the left-hand side showing the full raw distribution and the right-hand side showing the donut hole distribution that drops the students at the exact cut-off for their school and application year. As the figures make clear, dropping students with scores exactly at the cut-off does little to reduce the mass points that arise in the underlying distribution. The irregular distribution of raw test scores makes it impossible to conduct the standard McCrary tests for manipulation at the boundary. As a result, we conduct a series of simulations in order to provide confidence that the observed features of the data arise naturally from the admissions system.

2.1 The Effect of Dropping Students at Threshold

As shown in Figure 2 and Figures 1B-8B, the elimination of applications who have the same score as the admissions letter threshold selected for their school and year does little to change the irregular nature of the distribution. In order assess how much of the change in the distribution we can explain, we calculate the square root of the mean squared error deviation between the full empirical distribution and the distribution minus the applications at the cut-off for each specific application school and year. These calculations are conducted separately for the periods 06-08, 09-10 and 11-13 identified in Table 1 and shown in Figure 2. We then conduct a simple simulation that involves the following steps:

1) Select the cut-offs chosen by all schools as a population of relevant cut-offs for that year in order to take into account the fact that the rules change from year to year likely changing the location of any mass points.
2) For each school and year, select one cut-off randomly from the population of cut-offs for that year omitting the school's own cut-off for that year.
3) Pool the data across schools and relevant years to obtain a simulated distribution.
4) Delete all applications from the pool for that school and year at the simulated cut-off, and calculate the root mean squared error from the full simulated distribution.
5) Repeat steps 2 through 4 and average the resulting root mean squared errors.
6) Smooth the raw score distribution for each year over all schools using a five point moving average centered on the score for which the distribution is being calculated.
7) Repeat steps 2 through 5 using the probabilities from the smoothed distribution to select the score for the simulated cut-offs.

While the simulation cannot completely explain the deviations caused by dropping applications at the cut-off, the simulations explain much of the deviation especially for 06-08. Specifically, we calculate the difference between the mean of our simulated root mean squared error using simulated cut-offs drawn from the schools' cut-offs for the same year and the mean using the smoothed distribution. We then divide that difference by the difference between the empirical root mean squared error and the mean using the smoothed distribution. The smoothed distribution provides a baseline effect of the expected difference between the distributions after dropping applications at likely scores. The effect of dropping applications at one of the other school's cut-offs relative to that baseline explains 71 percent of the effect of dropping applications at the school's actual cut-off for 06-08. This exercise explains 62 and 60 percent of the effect for 09-10 and 11-13, respectively.

2.2 Simulation of Raw Test Score Distribution

We have argued that the discrete nature of the individual components in the total score and the natural correlation between these components leads to the large number of mass points and holes in the empirical application score distribution. To demonstrate that these components can generate these patterns, we next conduct the following exercise.

1) We measure the correlation between the four or six components in each year.
2) We use these correlations and draw simulated data for each of these components from a multivariate normal distribution in order to match the exact number of applications at each school and year.
3) We then map these components through the normal CDF to obtain a probability, and then assign discrete scores to each component based on the empirical frequency of each discrete score for that school and year.
4) We add the components, and the total resulting scores form the simulated distribution for each year.

Appendix Figures 1B-8B present the results of this exercise for each year between 2006 and 2013. As mentioned above, the top panel of Figures 1B-8B shows the raw distribution, and the bottom panel of each figure shows two separate simulations of the distribution. While we cannot perfectly replicate the mass points and holes in the true distribution (likely because we have to assume a form for the unknown, underlying latent distribution), these simulations consistently generate a substantial number of sizable mass points and holes in the simulated data. Note that in 2009 we do not observe the sub-scores separately for extracurricular activities and the written statement and so conduct the simulation in 2009 using just the original four components.

2.3 Simulation of Centered Test Score Distribution

Another unusual feature arises once we center the application score distributions by school and year. The left-hand side of Appendix Figures 9B-16B show the centered application score distribution by school and year and there is a substantial cliff to the left of the cut-off location in every case. These cliffs could be consistent with substantial manipulation, placing students to the right of the cut-off with some noise so that an unexplained mass of students is just to the right of the cut-off. However, it is also possible that the cliffs in the centered score distribution arise naturally from the mass points and holes in the raw distribution. If a school is trying to issue a specific number of acceptances and typically works down the list from high to low scores, then the stopping point for this process is more likely to land on a mass point

as opposed to a hole because the mass points contain many students and any of those students could help fill the school's capacity.

To examine whether we can replicate the cliffs we observe in the centered application score distributions, we conduct another simulation. For each school, we capture the number of students admitted each year, and view these numbers as an empirical distribution of the number of students the school might admit. For each school, we also calculate the fraction of students over all years who received an admission letter whose test score falls above the threshold, and the fraction of students over all years who are above the threshold, but did not receive an admissions letter as a share of the number of students above the threshold. We then use these ratios to scale the potential numbers of students admitted up, based on some students not being admitted even though they were above the threshold, and down based on students being admitted whose scores were below the threshold. We then conduct the following exercise:

1) For each year and school, we select a scaled number of admitted students from the empirical distribution over all years for that school.
2) We then count down the distribution from high to low test scores for the year and school until we have admitted the scaled number of students. The score of the last student admitted is the simulated threshold test score of cut-off.
3) We then use the simulated cut-offs for every year and school to center the test score distribution on the simulated cut-off and then pool the centered distributions over all schools.

The result of this exercise is shown on the right-hand side of Figures 9B-16B. Specifically, as noted above, for each year, the left-hand side of Figures 9B-16B show the actual empirical distribution of the re-centered application score and the right-hand side shows the simulated distribution based on the process described above. The resulting simulated distributions are relatively similar to the empirical distributions for the years of 2006-2009, but starting in 2010, and especially in 2011, the simulated centered distributions become much smoother while we continue to see mass points and cliffs in the empirical distribution. While pooling centered distributions should in principle smooth over mass points, we have no explanation for the time pattern of this phenomenon. However, given our inability to generate data that looks similar to the actual centered distribution after 2009, we rerun all our analyses dropping the years between 2010 and 2014. These results are shown in Appendix Table 2A and as noted in the paper, with the exception of college attendance, our core results are robust to this sample restriction.

Table 1B
Correlations between Estimated Mean Threshold Sizes

Correlations by School over Sample of Years			
Simulation #'s	1	2	3
2	0.871916		
3	0.773716	0.819969	
4	0.840485	0.912581	0.89444

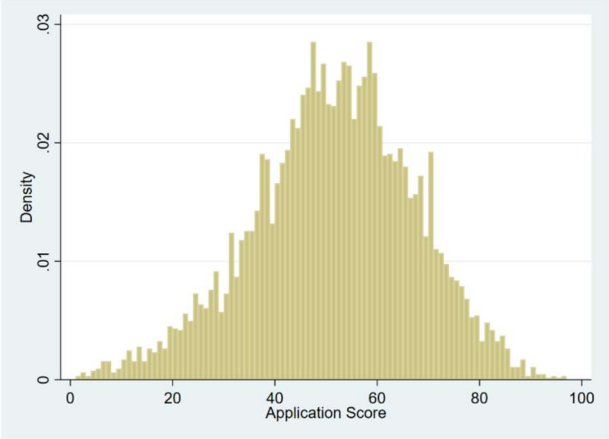
Correlations by Year over Sample of Schools			
Simulation #'s	1	2	3
2	0.136102		
3	0.557161	0.470346	
4	0.402145	0.051942	-0.10463

Table 2B
Estimated Thresholds for one Simulated Set of Hold-out Samples

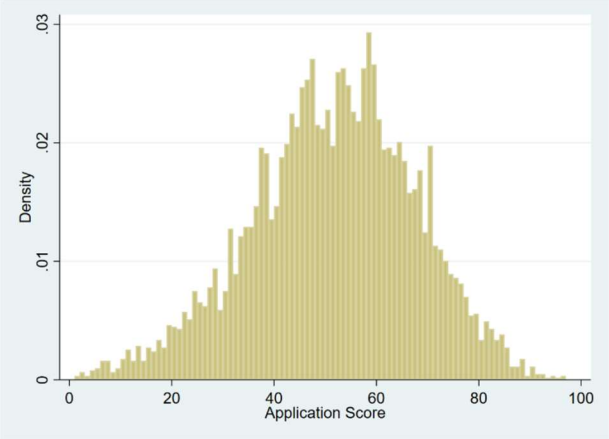
Application School/Year	2006	2007	2008	2009	2010	2011	2012	2013	Mean Threshold	Significant at 0.10
911	0.871	0.266	0.881	0.586	0.921	0.831	0.903	0.704	0.745	1.000
912	-0.053	0.411	0.432	0.499	0.450	0.426	0.558	0.652	0.422	0.750
913	0.333	0.493	0.529	1.116	0.374	1.021	-0.299	0.924	0.561	0.750
914	0.506	0.610	0.652	0.271	0.648	0.726	0.818	0.695	0.616	0.875
915	0.563	0.658	0.520	0.602	0.337	0.858	0.647	0.425	0.576	1.000
916	0.674	0.646	0.563	-0.001	0.172	0.228	0.073	0.605	0.370	0.625
917	0.706	0.693	0.100	0.791	0.805	0.581	NA	0.596	0.610	0.750
918	0.592	0.836	0.506	0.697	0.515	0.013	-0.729	0.266	0.337	0.750
919	0.375	0.745	0.821	0.795	0.594	0.582	0.556	0.461	0.616	1.000
920	0.747	0.297	0.341	0.138	0.726	0.461	0.109	0.475	0.412	0.750
922	0.763	0.756	0.359	0.901	-0.094	0.734	0.620	0.152	0.524	0.750
923	0.652	0.606	0.322	0.117	0.810	0.291	0.670	0.436	0.488	0.875
924	0.790	0.474	0.576	-0.692	0.728	0.559	0.560	-0.362	0.329	0.875
925	0.436	0.281	0.528	0.447	0.802	0.865	0.870	0.747	0.622	0.875
926	0.777	0.487	0.665	0.476	0.723	0.599	0.544	0.348	0.577	1.000
927	1.077	0.560	0.681	0.002	0.587	0.582	0.875	0.465	0.604	0.750
Mean Threshold	0.613	0.551	0.530	0.422	0.569	0.585	0.452	0.474	0.525	
Fraction Significant at 0.10	0.938	0.875	0.938	0.688	0.813	0.938	0.733	0.813		0.843

Figure 1B
Application Score Distributions: Application Year 2006

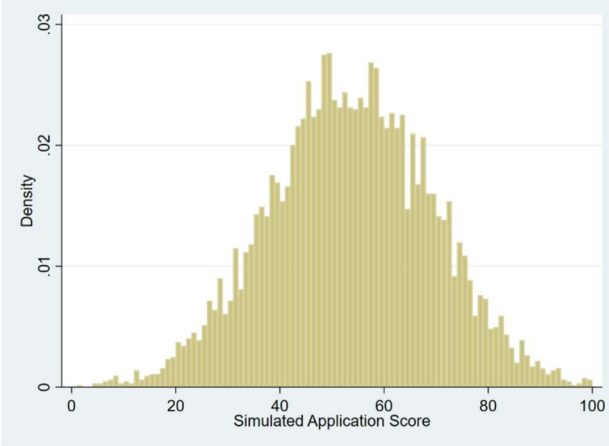
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

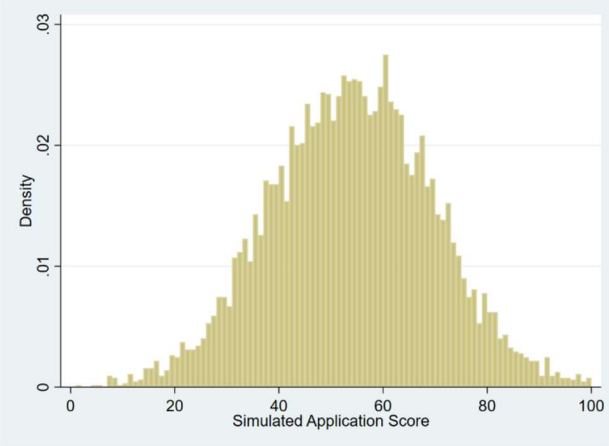
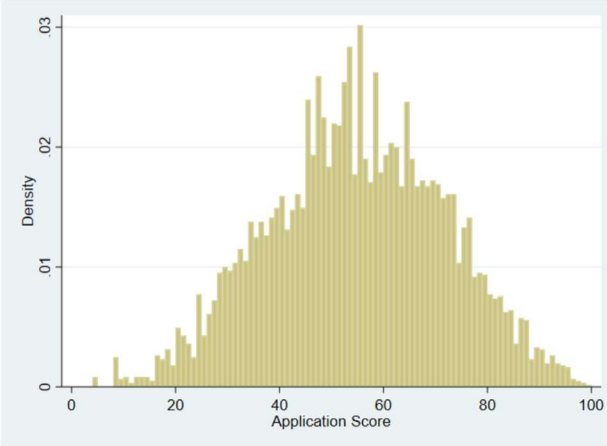
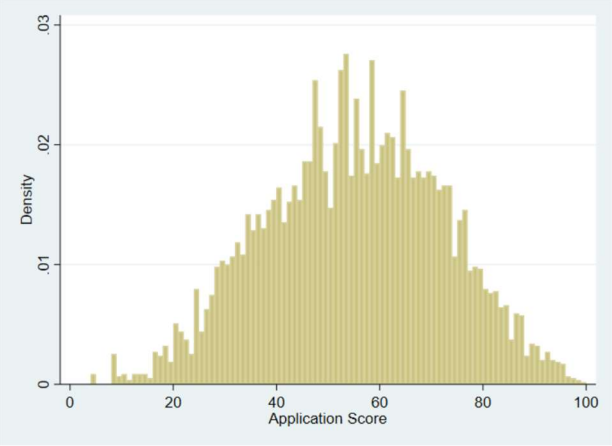


Figure 2B
Application Score Distributions: Application Year 2007

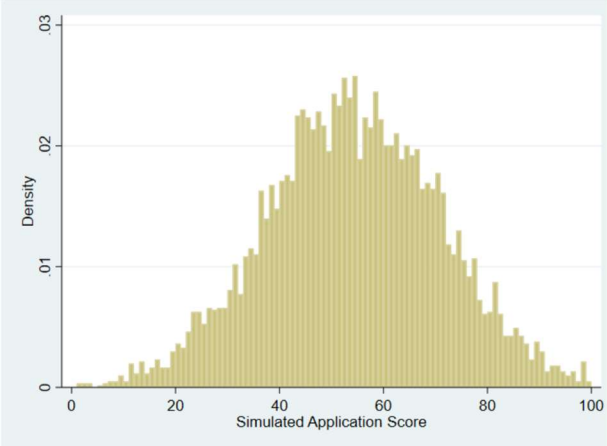
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

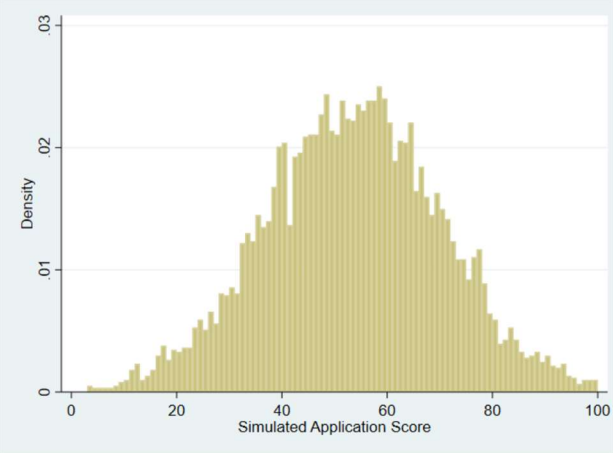
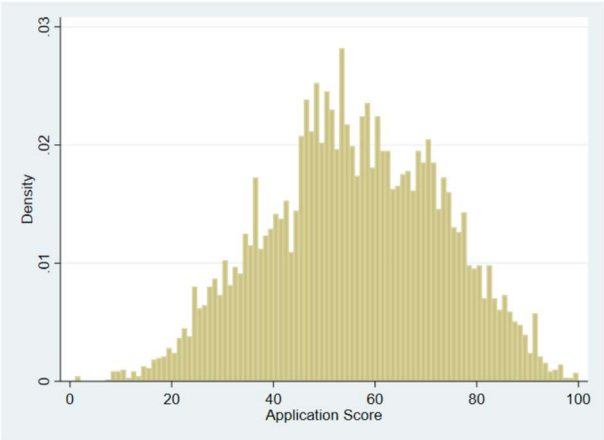
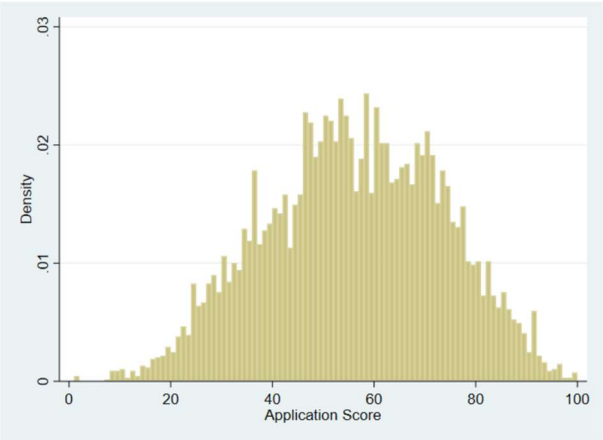


Figure 3B
Application Score Distributions: Application Year 2008

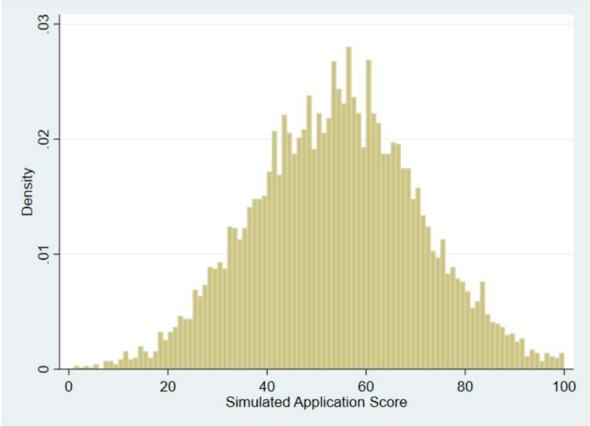
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

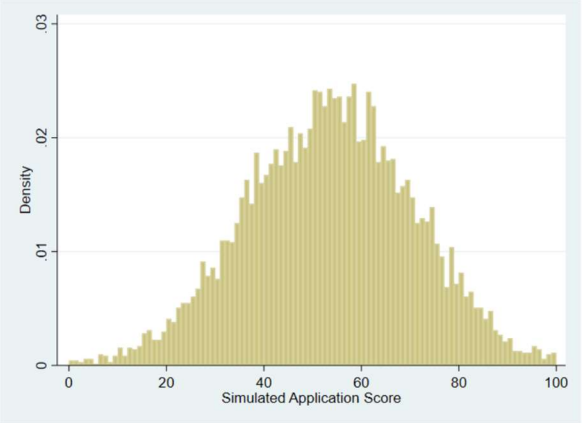
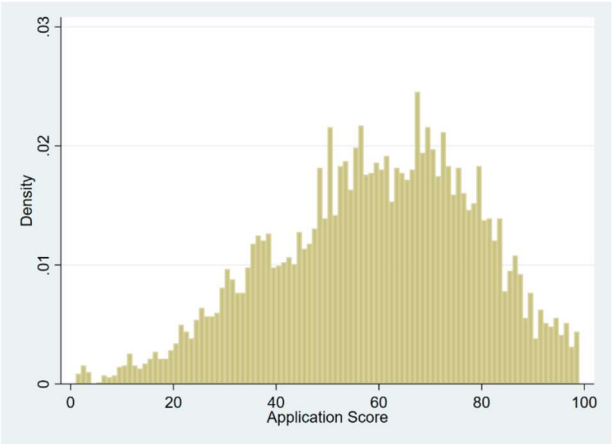
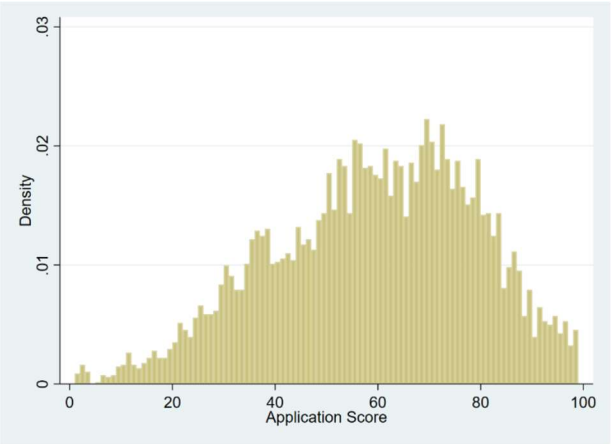


Figure 4B
Application Score Distributions: Application Year 2009

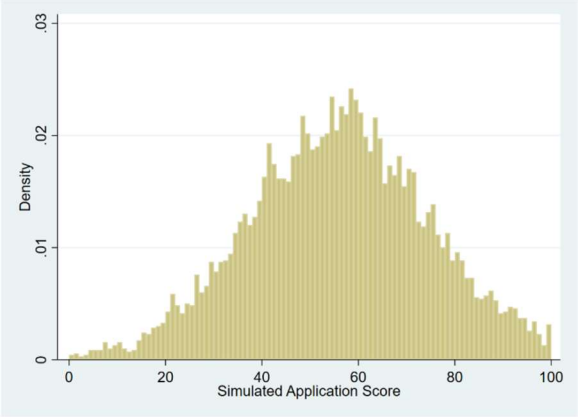
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

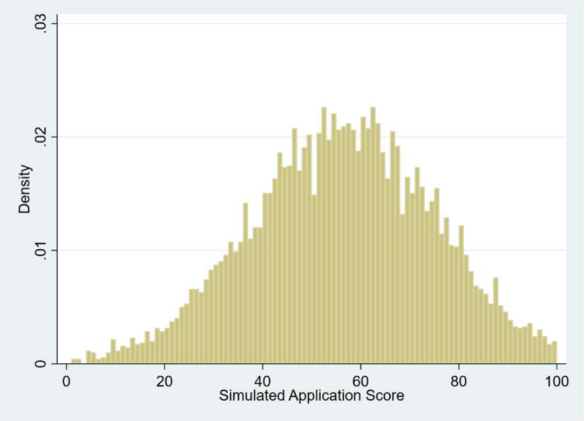
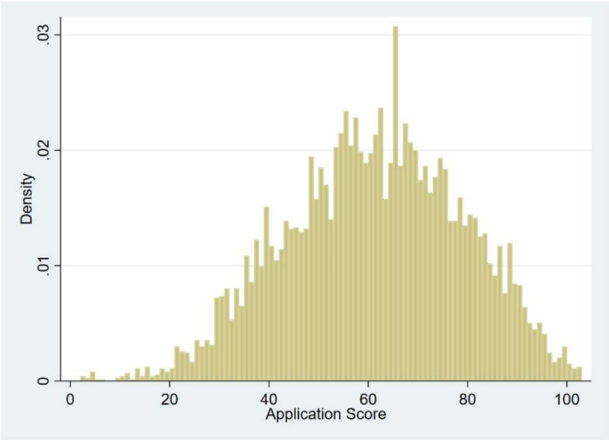
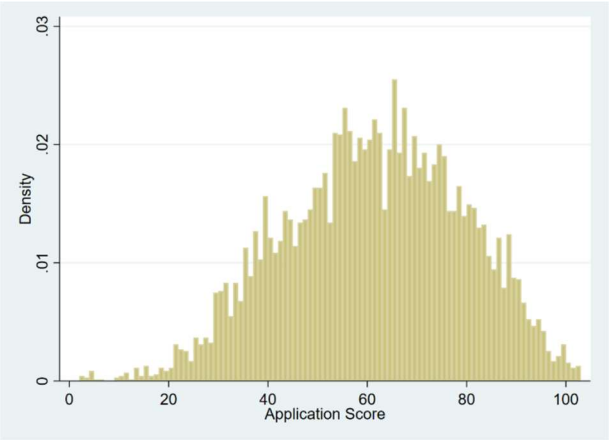


Figure 5B
Application Score Distributions: Application Year 2010

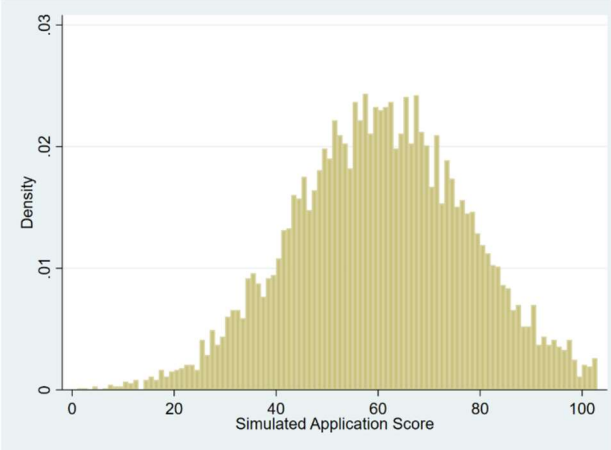
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

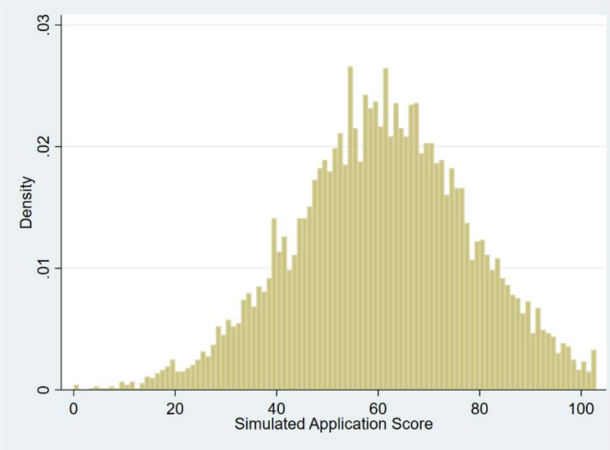
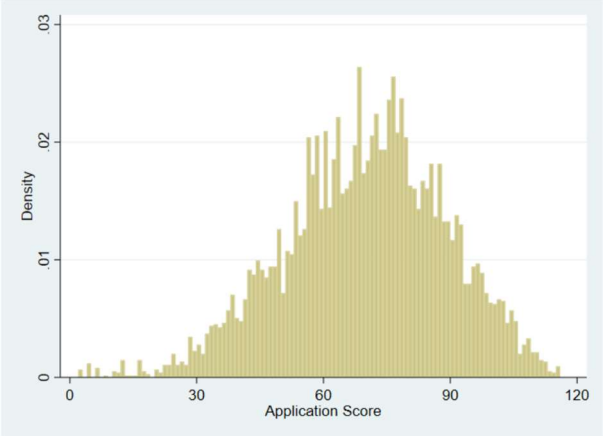
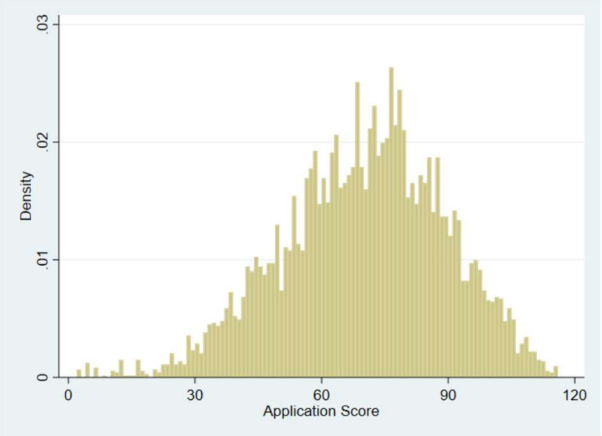


Figure 6B
Application Score Distributions: Application Year 2011

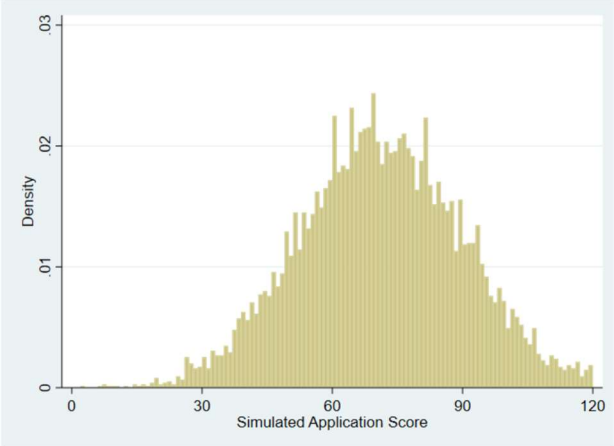
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

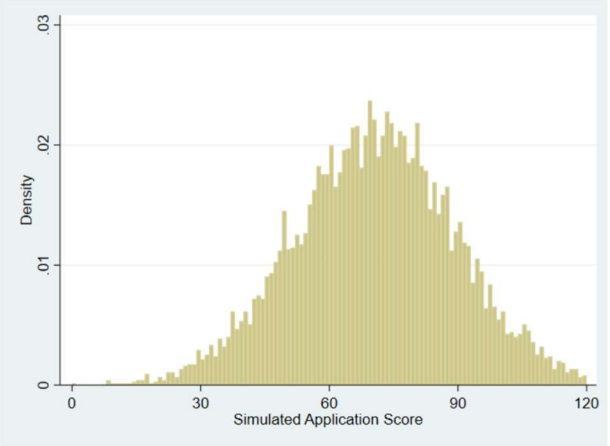
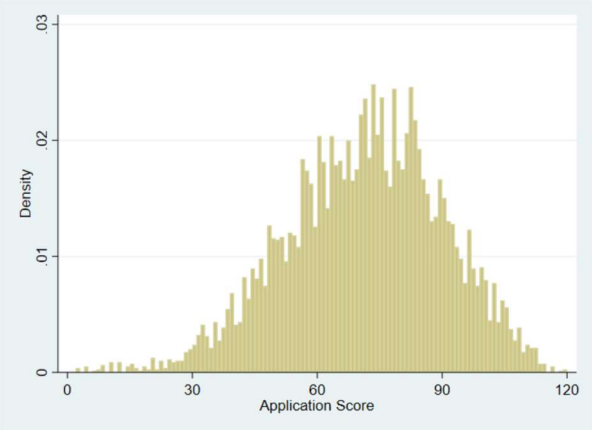
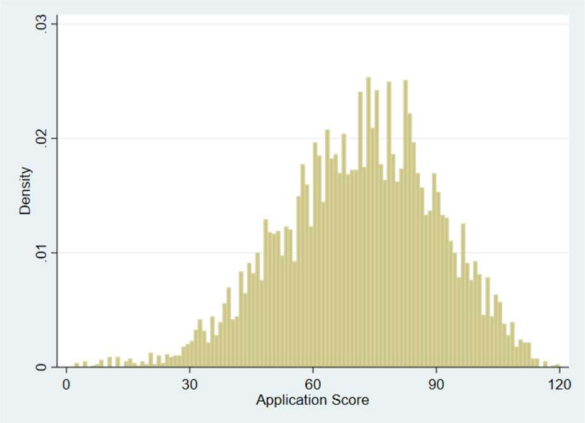


Figure 7B
Application Score Distributions: Application Year 2012

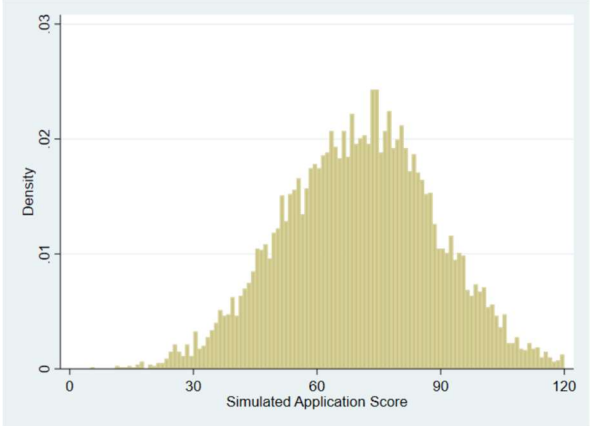
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

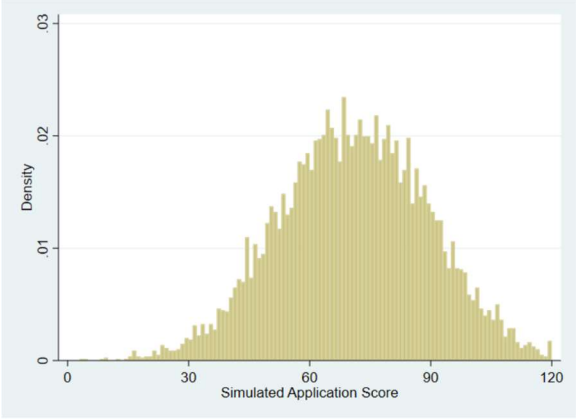
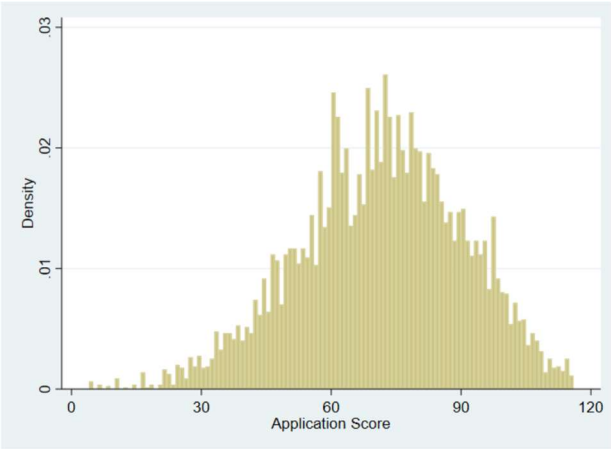
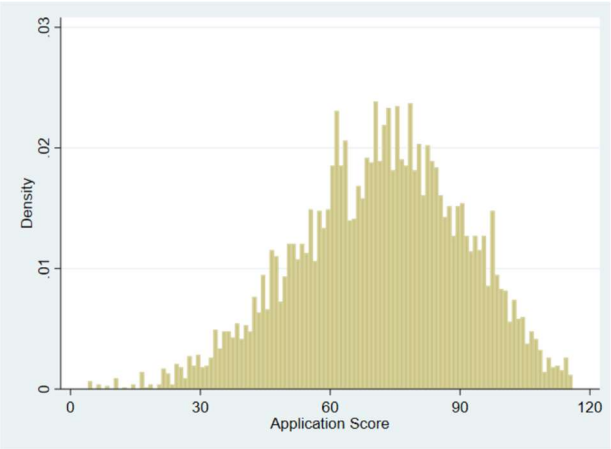


Figure 8B
Application Score Distributions: Application Year 2013

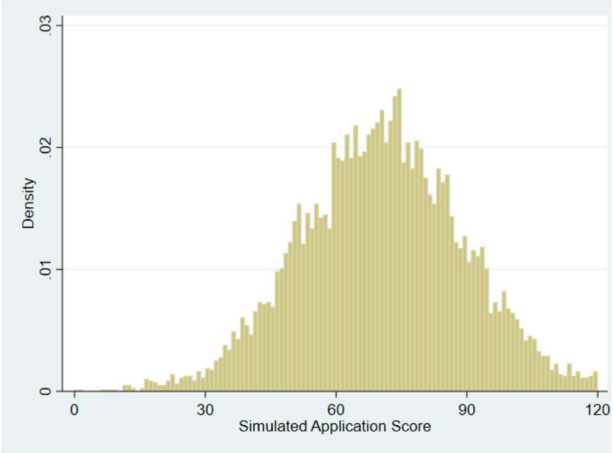
Full Sample



Donut Hole Sample



Simulated Full Sample



Simulated Donut Hole Sample

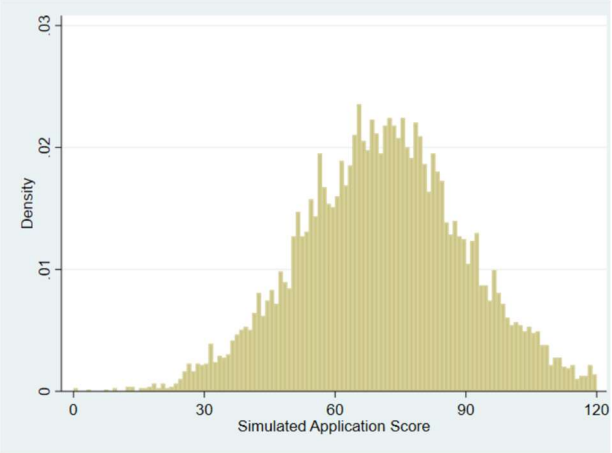
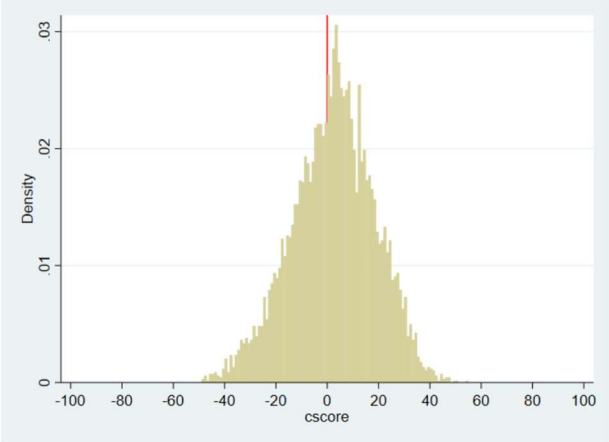


Figure 9B
Centered Score Distributions: Application Year 2006

Full Sample



Simulated Sample

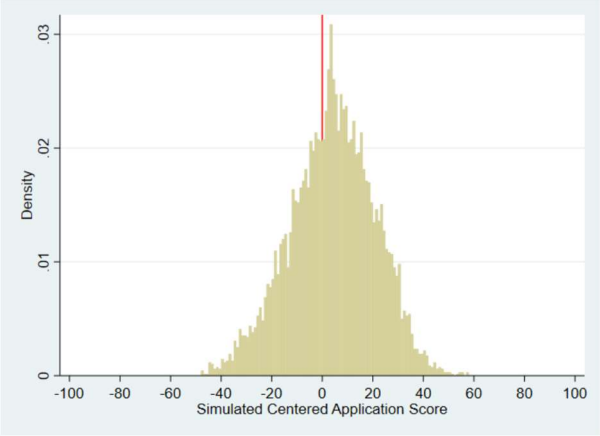
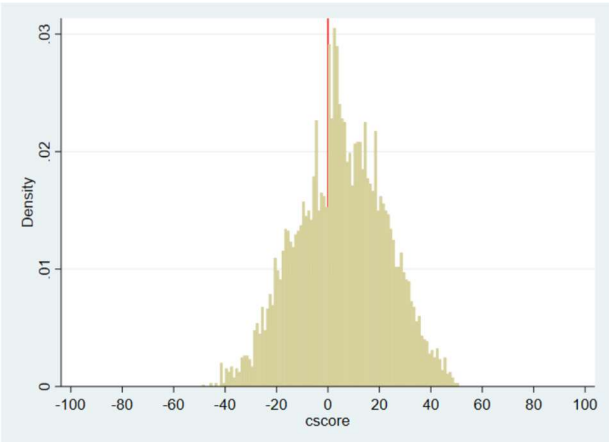


Figure 10B
Centered Score Distributions: Application Year 2007

Full Sample



Simulated Sample

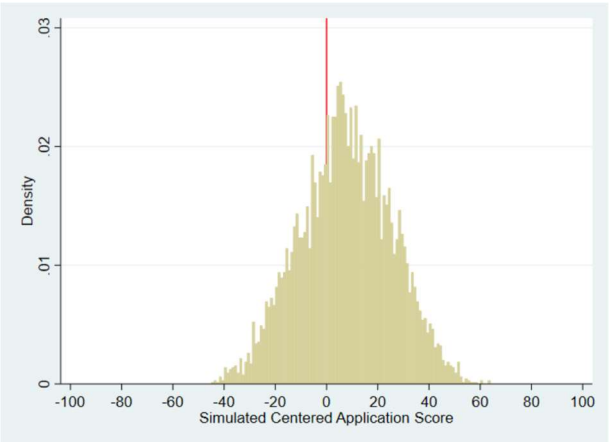
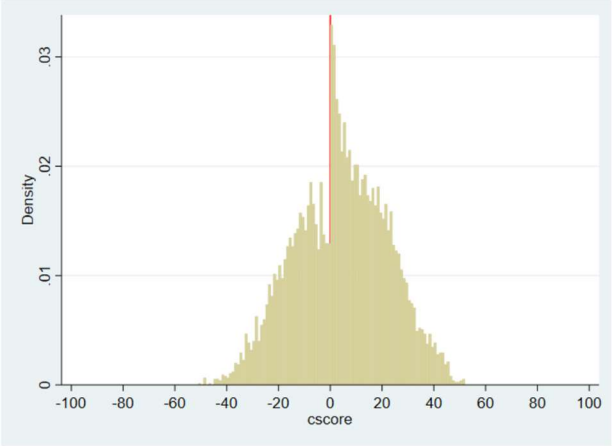


Figure 11B
Centered Score Distributions: Application Year 2008

Full Sample



Simulated Sample

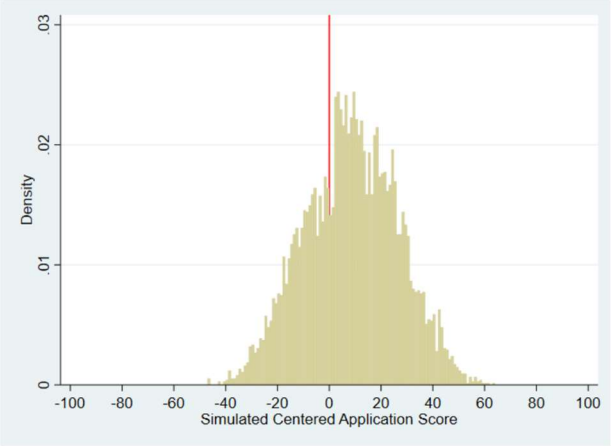
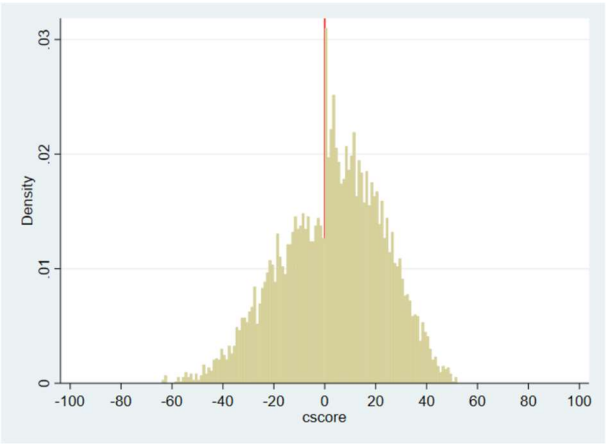


Figure 12B
Centered Score Distributions: Application Year 2009

Full Sample



Simulated Sample

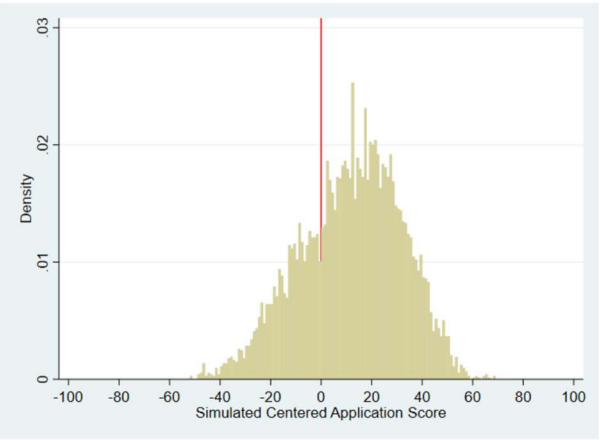
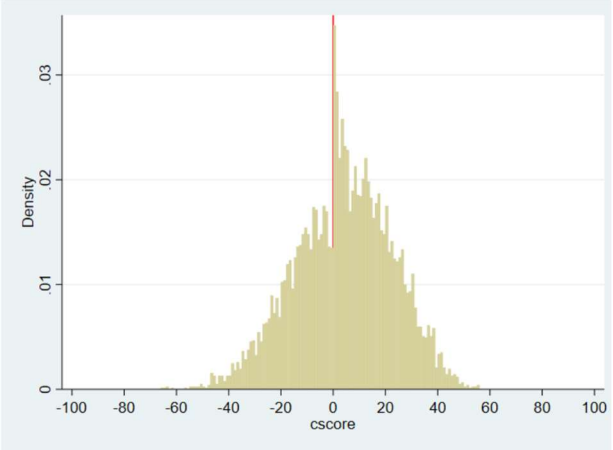


Figure 13B
Centered Score Distributions: Application Year 2010

Full Sample



Simulated Sample

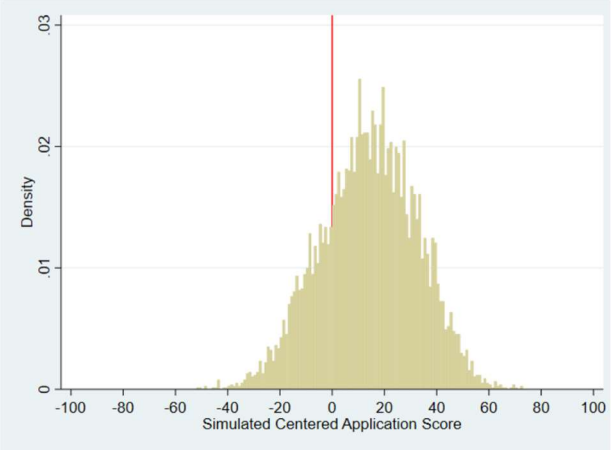
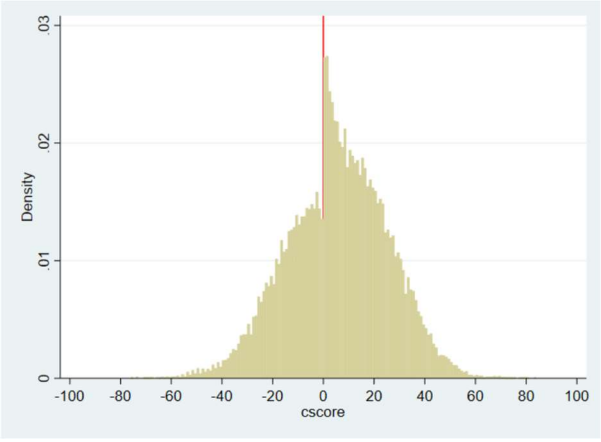


Figure 14B
Centered Score Distributions: Application Year 2011

Full Sample



Simulated Sample

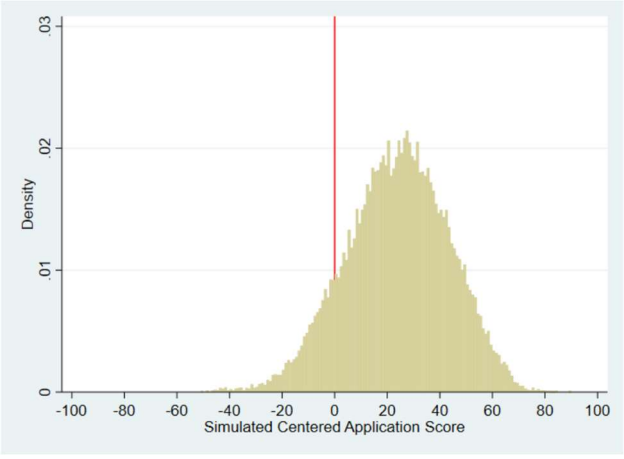
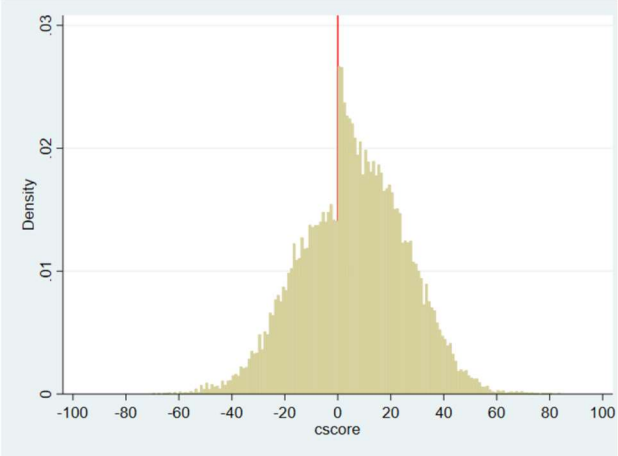


Figure 15B
Centered Score Distributions: Application Year 2012

Full Sample



Simulated Sample

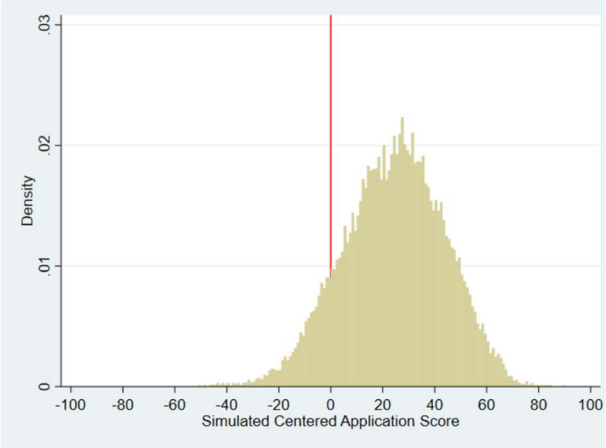
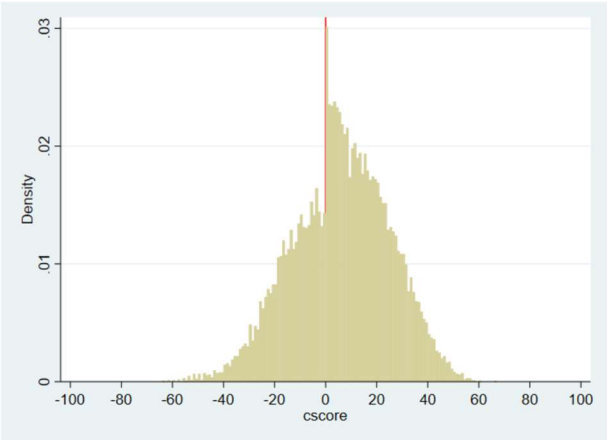


Figure 16B
Centered Score Distributions: Application Year 2013

Full Sample



Simulated Sample

