ESSAYS ON ECONOMICS OF EDUCATION

by

Minseon Park

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2023

Date of final oral examination: 05/08/23

The dissertation is approved by the following members of the Final Oral Committee:
John Kennan, Professor, Economics
Christopher Taber, Professor, Economics
Chao Fu, Professor, Economics
Jesse Gregory, Associate Professor, Economics
Alina Arefeva, Assistant Professor, Business

© Copyright by Minseon Park 2023
All Rights Reserved

To my grandma

ACKNOWLEDGMENTS

My first thanks goes to my main advisor, John Kennan. He showed endless patience and support during the entire process. I also appreciate all the help from other committee members, Chris Taber, Chao Fu, and Jesse Gregory. I learned immensely from numerous interactions with my advisors; the contrast between the cluelessness from the meetings in the early years and the intellectual thrill from the meetings in the last year assured me that I grew a lot as a researcher. Without all their feedback and help, I would not have been able to produce the chapters in this thesis. Moreover, they have been and will forever be my role models to be a serious, hardworking, communicative, down-to-earth, and passionate researcher.

I learned a lot from the whole intellectual community at the department beyond my advisors. I thank Jean-Francois Houde for always giving knowledgeable answers to my countless questions. I thank Jeff Smith for all the follow-up notes from my presentations and for pushing me to be an institutional-background expert. I thank Matt Wiswall, Naoki Aizawa, Corina Mommaerts, Rasmus Lentz, and Panle Jia Barwick for constructive feedback on my work. I also thank Job Boerma for sharing his fresh knowledge on how to navigate the intense job interview and visit process.

I am deeply indebted to my coauthors of the chapters of this thesis. I thank Dong Woo Hahm for introducing me to the world of school choice. I learned from his rigorousness and efficiency as a researcher. I also thank Lois Miller for being an understanding coauthor, English teacher, and personal higher-education-literature librarian. I cannot imagine a better match in terms of both complementarities in skills and respect for each other.

I was very fortunate to be surrounded by a supportive and efficient department's staff team, especially Kim Grocholski, Becca George, Vicki Fugate, and Julie Anderson. They have answered my numerous questions and understood my forgetfulness. I also thank our previous chair, Ananth Seshadri, for his tireless work in providing research infrastructure. All the funding and awards I have received from the Department of Economics have not only financially supported me but also given signals that I was on track.

I could improve all three chapters thanks to feedback from researchers in the field — Sandra E. Black, Michael Best, Yeon-Koo Che, Pierre-Andre Chiappori, Gautam Gowrisankaran, YingHua He, Seungki Hong, Simon Lee, W. Bentley MacLeod, Bernard Salani´e, Suk Joon Son, Olivier Tercieux, Miguel Urquiola, Milena Almagro, Rajeev Darolia, Margaux Luflade, Doug Webber, Drew Anderson, Emily Cook, Nick Hillman, Sarah Pingel, Andria Smythe. We would also like to thank the seminar participants at Columbia University, the University of Wisconsin-Madison, and the Federal Bank of Richmond. Thanks also goes to the New York City Department of Education, particularly to Benjamin Cosman, Broggini Matthew, Stewart Wade, and Lianna Wright.

My accomplishments would not have been possible without fellow students. I have been lucky to have my office mates — Zau Aitkulova, Jason Choi, and Lois Miller — with whom I could share lunch boxes, joys, tears, and much other stuff. I thank Natalia Serna Borrero for being such a cordial pacemaker. I thank Sarah Waldfogel for being a confidence booster. I thank every single one in my cohort and many lower-year students who shared time and energy to listen to my research ideas, progress, and presentations. I also thank senior students including Amrita Kulka, Joanna Venator, Anson Zhou, Sandra Spirovska, Arpita Patnaik, Siha Lee, and Gueyon Kim for sharing their time, knowledge, and materials. Thanks to them, we younger students could learn in the cordial research community amongst students and will pass that culture on. Special thanks goes to Annie Soyean Lee for being a sincere friend, mentor, and unconditional caregiver.

Lastly, I want to thank my family and Andrew Bonjune Choi. To Mom and Dad– thank you for all your support. To my sister, thank you for all your support and sorry for not being able to be physically beside you whenever you want. To Andy, you are truly a hidden coauthor of all the chapters. I am indebted to all the sharp questions from you and serious discussions with you. I do not think I was able to reach the finish line safely without your emotional and financial support.

All errors are mine.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation includes three essays in the field of economics of education.

The first chapter (joint with Dong Woo Hahm) explores the impact of public school assignment reforms by building a households' school choice model with two key features—(1) endogenous residential location choice and (2) opt-out to outside schooling options. Households decide where to live taking into account that locations determine access to schools—admissions probabilities and commuting distances to schools. Households are heterogeneous both in observed and unobserved characteristics. We estimate the model using administrative data from New York City's middle school choice system. Variation from a boundary discontinuity design separately identifies access-to-school preferences from other location amenities. Residential sorting based on access-to-school preference explains 30% of the gap in test scores of schools attended by minority students versus their peers. If households' residential locations were fixed, a reform that introduces purely lottery-based admissions to schools in lower- and mid-Manhattan would reduce the cross-racial gap by 7%. However, households' endogenous location choices dampen the effect by half.

The second chapter (joint with Dong Woo Hahm) explores how students' previously attended schools influence their subsequent school choices and how this relationship affects school segregation. Using administrative data from New York City, we document the causal effects of the middle school a student attends on her high school application/assignment. Motivated by this finding, we estimate a dynamic model of middle and high school choices. We find that the middle schools' effects mainly operate by changing how students rank high schools rather than how high schools rank their applications. Counterfactual analysis shows that policymakers can design more effective policies by exploiting the dynamic relationship of school choices.

The third chapter (joint with Lois Miller) studies how colleges' "sticker price" and institutional financial aid change during and after tuition caps and freezes using a modified event study design. While tuition regulations lower sticker prices,

colleges recoup losses by lowering financial aid or rapidly increasing tuition after regulations end. At four-year colleges, regulations lower sticker price by 6.3 percentage points while simultaneously reducing aid by nearly twice as much (11.3 percentage points). At two-year colleges, while regulations lower tuition by 9.3 percentage points, the effect disappears within three years of the end of the regulation. Changes in net tuition vary widely; focusing on four-year colleges, while some students receive discounts up to 5.9 percentage points, others pay 3.8 percentage points more than they would have without these regulations. Students who receive financial aid, enter college right after the regulation is lifted, or attend colleges that are more dependent on tuition benefit less.

1.1 Introduction

Across the U.S. in 2016, 33% of K-12 students lived in a school district where they could choose, to some degree, which public school to attend.[1] Public school choice systems provide students with multiple options beyond the nearest school to their home, standing in contrast to classic settings where students are automatically assigned to schools based on their home addresses. Such contrast has raised hope that centralized school assignments could decouple educational disparities from spatial disparities at scale. However, many popular schools under a choice system give admissions priority to students from residential locations nearby, even when they accept applications from a broader set of students (Dur et al., 2013). Such location-based admissions rules have triggered debate over the design of admissions rules, motivated by a concern that these contribute to the continued school segregation observed in many school choice settings (Cohen, 2021).

How effectively can we desegregate schools with reforms on the location-based admissions rules in a public school choice system? We answer this question by developing a households' school choice model that considers two important margins through which households may respond: residential location choice and opt-out to outside schooling options.

The key feature of the model is households' endogenous location choice. While previous work has documented that residential location explains half the racial gap in test scores of schools attended by students under centralized school choice (Laverde, 2020), *how* households make the residential location decision has received little attention in the school choice literature (e.g., Abdulkadiroğlu et al., 2017). In our model, households choose residential locations by considering *access to schools*, which refers to both admissions probabilities and commuting distances to schools that vary across locations.

[1] Based on authors' calculation using The National Center for Education Statistics 2019 National Household Education Surveys: Parent and Family Involvement in Education Survey.

We set up a multi-stage discrete choice model where households sequentially choose (1) which location to live in, (2) which school to apply to, and (3) whether to enroll in the assigned school or opt out to outside schooling options. Households have observed and unobserved heterogeneous preferences over a set of location and school characteristics, which leads to rich residential and school sorting patterns.

We start by providing causal evidence that households consider location-based admissions rules when deciding where to live. Our empirical context is the middle school choice system in New York City (NYC), where 70,000 students and 700 middle schools are matched each year. Each student has over 30 public school options to apply to, and which Community School District (CSD)—a subdivision of the city—they reside in largely determines the choice set and admissions probabilities. Leveraging this institutional aspect, we apply a boundary discontinuity design (BDD) to compare Census blocks that are close to one another but located on opposite sides of a CSD boundary. By doing so, we deal with the endogeneity concern that locations with higher admissions chances to high-achieving schools might have amenities unobserved to researchers but are observed and valued by households. Estimates indicate that Census blocks within a CSD with one standard deviation higher school test scores have 22% more households with middle school applicants.

We use the variation from the BDD to identify how much households value access to school relative to other location amenities. We estimate our structural model using an extension of the expectation-maximization algorithm with a sequential maximization step (ESM, Arcidiacono and Jones, 2003). This keeps the estimation tractable while enabling us to jointly estimate all stages of the model to account for households' selection into locations.

The results show that endogenizing households' residential choice has important implications for (1) understanding the source of school segregation under the status quo, (2) obtaining an unbiased commuting cost estimate, and (3) predicting the implications of the counterfactual policy.

First, our estimates illustrate that households' location choices based on location's access to schools play a large role in explaining which students are matched to which schools. To show this, we shut down each part of the model in a decomposition exercise. 30% of the gap in test scores of schools attended by minorities versus non-minorities is explained by households' residential sorting based on locations' access to schools. Households' heterogeneous preference over other location characteristics and school characteristics explains 45% and 18% of the cross-racial gap, respectively.

Second, we find that a model that does not account for endogenous location choice overestimates commuting costs by 15%. Our model estimates show that a median household is willing to pay $19 per school day to reduce commuting time to school by 50 minutes. Commuting cost is an important parameter that governs the degree to which students take advantage of school choice options rather than applying to schools nearest to their residential locations. The reason for the overestimation is that households choose locations near schools they prefer since it increases their admissions probabilities. This leads to a spurious result in which they apply to schools nearby not because they care about distance but because in their location choice they cared about admission probability. Without correcting for households' selection into locations, the model would misinterpret households' applying to schools nearby as solely due to commuting costs. Due to residential sorting based on unobserved school preference, this is still true when one controls for households' observed characteristics.[2]

Finally, we describe how households' spatial reshuffling in response to a school desegregation reform can affect the effectiveness of the policy. We consider a counterfactual policy that introduces purely lottery-based admissions to schools in District 2, the district with the highest test scores and housing costs. Covering lower- and mid-Manhattan, District 2 has been at the center of ongoing policy debates regarding the design of location-based admissions criteria.[3] When we fix households' residential locations, lottery-based admissions to District 2 schools

[2]For example, a household that puts a higher value on school safety than other observably similar households will sort into locations that increase their child's admission chances into a safer school.

[3]Shapiro, Eliza, N.Y.C. to Change Many Selective Schools to Address Segregation, the *New York Times*, December 18, 2020.

would close the cross-racial gap in school test scores by 7%. This is because some minority students residing outside the district are assigned to District 2 schools, which pushes out non-minority District 2 residents to lower-achieving schools.

However, households' location choices in response to the policy dampen the equity impact by half. Two types of spatial reshuffling exert opposing forces. On the one hand, some minority households choose residential locations closer to District 2 in response to the reform. With shorter commuting distances to District 2, they are more likely to apply to District 2 schools. Spatial reshuffling of this sort amplifies the desegregation effect of the policy. On the other hand, most of the non-minority households who reside in District 2 under the status quo relocate out of the district. Since other districts still have location-based admissions in place, they seek other locations that assure higher admissions probabilities to high-achieving schools. Such spatial reshuffling dampens the equity effect of the policy.

The equilibrium force amplifies the second reshuffling while muting the first. This is because purely lottery-based admissions to District 2 schools induce more applications, and thus the equilibrium admissions cutoffs of these schools increase. This weakens the incentive of minority households to relocate closer to District 2 but strengthens that of non-minority households to relocate farther from District 2. We find that households substitute between opting-out to outside schooling options and choosing different residential locations. But, overall, opt-out plays a smaller role in determining the effectiveness of the reform on reducing the cross-racial gap.

Related Literature We contribute to two strands of the literature. First, we extend the school choice literature by considering households' endogenous location choice. While it is well known that residential location is the main source of school segregation (Laverde, 2020), little is known about *how* households choose where to live in response to the design of centralized school choice. Previous studies have focused on assignment mechanisms (Abdulkadiroğlu et al., 2015a, 2017; He, 2015; Agarwal and Somaini, 2018a; Che and Tercieux, 2019a; Calsamiglia et al., 2020); information provision (Hastings and Weinstein, 2008a; Hoxby and Turner, 2015a; Luflade, 2018; Corcoran et al., 2018b; Chen and He, 2021b; Fack et al., 2019a; Allende et al., 2019); limited attention (Ajayi and Sidibe, 2020; Son, 2020); and previously attended schools (Hahm and Park, 2022).

By modeling households' endogenous location choices, we first compare the implications of counterfactual policies when households' residential locations are fixed versus adjusted. This approach aligns with reduced-form evidence that access to school shapes the composition of residents and housing costs of locations (Black, 1999; Reback, 2005; Brunner et al., 2012; Schwartz et al., 2014; Billings et al., 2018). Moreover, we correct selection into locations in estimating school preference to obtain an unbiased estimate of commuting costs; we take a departure from the standard assumption in the literature that distances to schools are uncorrelated with households' unobserved school tastes conditional on their observable characteristics.[4]

Second, this paper adds to a large body of studies on within-city residential sorting, by studying households' location choice in a newly relevant setting of centralized school assignments. Among many papers in this literature, more closely related are those that give special attention to schools compared with other location amenities.[5] Earlier studies have focused on classical settings where each residential location is zoned to one public school (Bayer et al., 2007) while incorporating limited forms of school choices such as private school vouchers or inter-district transfers (Manski, 1992; Nechyba, 2000; Epple and Romano, 2003; Ferreyra, 2007).

Under centralized school assignments, households choose among many public schools from a given location. This enables us to study households' heterogeneous values over a set of school characteristics, including commuting distance. Indeed, this two-way heterogeneity is one of the main sources of school segregation under school choice settings (Idoux, 2022; Hahm and Park, 2022, e.g.). In contrast, frameworks in the location choice literature (Bayer et al., 2007) have considered

[4]This assumption is often found in the broader economics of education literature, which uses distance to schools as an instrumental variable for school application and attendance (Card, 1993; Schwartz et al., 2013; Walters, 2018; Mountjoy, 2022)

[5]Broader set of papers have studied how residential sorting is determined by other factors such as access to work (Ahlfeldt et al., 2015), ease of commuting (Barwick et al., 2021), consumption amenities (Almagro and Domınguez-Iino, 2019; Miyauchi et al., 2022), or neighborhood composition (Davis et al., 2019).

one-dimensional school characteristics, usually mean test scores, due to lack of variation coming from their setting where each location is zoned to one public school.

With the recent popularity of centralized school assignments, there have been a few papers proposing a unified framework of location choice and school choice. These include theoretical models (Xu, 2019; Avery and Pathak, 2021; Grigoryan, 2021) and a quantitative model (Agostinelli et al., 2021). Our paper complements theoretical models by estimating our model using data.

The closest paper to ours is by Agostinelli et al. (2021), from which we differentiate in two respects. First, our model features rich heterogeneity in households' location and school preferences. For example, Grigoryan (2021) shows that preference heterogeneity is crucial in determining the welfare implication of a school choice design.[6] We depart from the assumption that households have the same ordinal preferences over schools. We also consider location sorting based on unobserved school preferences to obtain unbiased commuting costs. Second, we model outside schooling options, another margin that some households use with the introduction of a more extensive school choice system.

Organization The remainder of the paper is organized as follows. Section 1.2 describes the public middle school choice system in NYC and the data. Section 1.3 presents motivating evidence on the interaction between residential location choice and school choice. Section 1.4 describes the model. Section 1.5 describes the empirical strategy and presents estimation results. Section 1.6 investigates the source of school segregation. Section 1.7 studies the equity impacts of a school desegregation reform.

[6]See Almagro and Domınguez-Iino (2019) for a similar discussion in a model where households have heterogeneous preferences over a set of urban amenities.

1.2 Institutional Background and Data

1.2.1 Public Middle School Choice in NYC

Each year, about 70,000 entering students and 700 middle school programs participate in the NYC citywide middle school choice system. There are about 500 middle schools. Multiple programs with separate curriculum can be offered by one school, and students apply to each program. In the following, we use the terms "program" and "school" interchangeably when there is no confusion. Schools that are part of the centralized choice system are governed by the city. The property tax rate is constant within the city and the city allocates the pooled funding to schools directly, largely based on the number of students.[7]

The main round of the school choice process starts in December of students' last year of elementary school. Students are given a customized list of programs they are eligible for and submit a rank-ordered list (henceforth, ROL) by designating their preference rankings over schools. In 2014-15, the average student had about 30 choice options. There is no list-length restriction, and students can list as many schools as they like (an example of an ROL is in subsection 1.9.1). The city uses the student-proposing deferred acceptance (SPDA) algorithm, which takes students' applications, schools' ranking over students, and the number of seats as main inputs and produces at most one assignment for each student (Gale and Shapley, 1962).[8]

Schools rank students by pre-announced admission rules, which consist of three layers. The first is eligibility, which determines students' school choice sets. If a student is not eligible for a program, she is never considered by the program, even when there are remaining seats. Second, eligible applicants are classified into a small number of priority groups. A program considers all students in the

[7] In 2002, Chapter 91 (Bill A.11627/S.7456-B) was enacted to reorganize the education system and has established centralized power. Since then, the public education system has been governed by the Panel for Educational Policy (PEP), which has 15 members; 9 of which are nominated by the mayor. The citywide school choice system was introduced in 2004 as part of this effort (Abdulkadiroğlu et al., 2005).

[8] 90% of students are assigned to a program on their list. The rest are matched to their fall-back option, which is usually a school in their attendance zone. See subsection 1.9.1 for details on the timeline and SPDA.

higher-priority group before considering any student in a lower priority group. Henceforth, we use the term "priorities" to refer to eligibility *and* priority groups, if not specified. Lastly, tie-breaking rules determine which students to admit among applicants of the same priority group. Some programs use a *nonrandom tie-breaker*, which is a school-specific function of the student's previous year's GPA, statewide standardized test scores, and punctuality. The rest use a *lottery* system in which each student receives one lottery number that applies to all such programs. See subsection 1.9.1 for more details on the admissions rules.

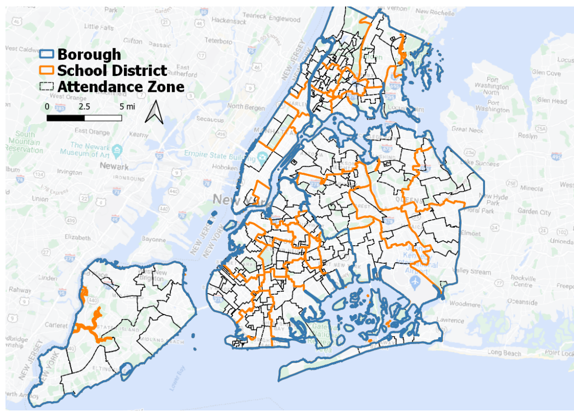


Figure 1.1: Geographic Divisions

Note: The city is divided into five boroughs (=counties), which are further divided into 32 school districts and 300 middle school attendance zones.

Students' residential locations are the main criterion for the eligibility and priority of schools. Figure 1.1 depicts different levels of geographic subdivisions that determine location-based admissions rules. The city is split into 5 boroughs, 32 Community School Districts (districts, henceforth), and more than 300 attendance zones.

Depending on their eligibility criteria, middle schools are classified into zoned programs, district programs, borough programs, and citywide programs. A student's residence or the location of her elementary school decides her eligibility for each type of school.[9] Of 669 programs in academic year 2014-2015, 14 were citywide

[9]87% of 2014-15 middle school applicants attended an elementary school in their residential

programs, 27 were borough programs, 478 were district programs, and the rest (150) were zoned programs. Schools can further assign priority based on finer geographic divisions. For instance, 81 of 478 district programs gave top priority to students from a particular attendance zone.

1.2.2 Data

Student and School Data Student-level data from the NYC Department of Education (DOE) cover middle school applicants in academic year 2014-15. The data have two crucial components for the purpose of this paper—students' school applications and residential Census block. The data also contain students' enrollment decisions, demographic characteristics, and statewide standardized test scores.[10]

We construct school characteristics by digitizing the *Directory of Public Middle Schools*.[11] It covers each program's admissions criteria, address, performance measures, previous year's capacity, and number of applicants. Students, parents, and guidance counselors use this as their primary information source during the middle school application process (Sattin-Bajaj et al., 2018). We augment this data by adding the number of crime incidents of different categories in each school building from a NYC Police Department's School Safety Report.

Housing Cost and Structure Housing cost and housing characteristics are from the NYC Department of Finance's (DOF) Rolling Sales files. The data include the exact address of each sold property, which is granular enough for us to observe on which side of a school district boundary the property is located. We describe the cleaning process of the DOF Rolling Sales files in detail in subsection 1.9.2.

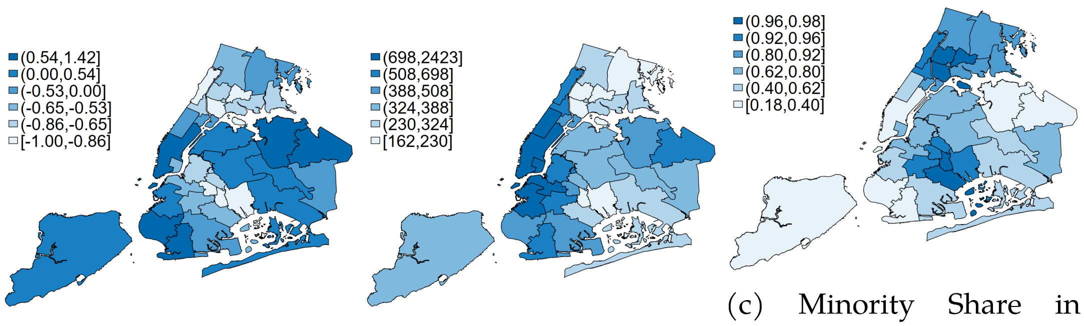
Amenities of Residential Location We construct location amenities from various sources. Land use comes from the Primary Land Use Tax Lot Output. We also

district.

[10]We focus on academic year 2014-15 because students can list only up to 12 middle schools in more recent years. With this list length restriction, students have the incentive to list less preferred schools with higher admissions chances (see Section 1.4).

[11]The city began publishing a digitized version in academic year 2017-18.

obtain consumption amenities such as the number of cafes from business licenses published by NYC Consumer and Worker Protection. Next, we collect information on bus stops, metro stations, and park areas using NYC OpenData GIS data files. We aggregate variables to Census block level. Finally, the demographic composition of each Census block group, such as ethnicity, age, education, and income, comes from the American Community Survey (ACS) 5-year estimates.



(a) School Test Score (b) Sales Price ($1,000) (c) Minority Share in Schools

Figure 1.2: Main Variables by District

Note: In panels (a) and (c), we take the average of the variables across schools within each district. The school test score is the average NYS standardized test scores of enrolled students. In panel (b), we present the average unit sales price of residential properties in each district.

Figure 1.2 presents the average characteristics for each district, which demonstrates a strong correlation among school achievement, housing cost, and share of minorities in schools. Summary statistics of main variables are in Table 1.8.

1.3 Motivating Data Pattern

1.3.1 Effect of Admissions Probability on Residential Sorting

This section presents evidence that households choose where to live by considering location-based admissions probabilities. Specifically, we show that locations with higher admissions chances to high-achieving schools have greater number of households with middle school applicants and higher housing costs. This makes the main motivation to model endogenous location choices under centralized school choice. Moreover, we show that these locations also have a lower minority share among households with middle school applicants, which implies that households have heterogeneous rates of substitution between housing cost and higher admissions chances to high-achieving schools. The main challenge to credibly show these patterns is that locations with higher admissions chances to high-achieving schools may have amenities unobserved to the econometrician but observed by and desirable to households, such as a well-kept playground.

To this end, we adopt a boundary discontinuity design (BDD) (Black, 1999; Bayer et al., 2007). Ideally, we would compare two locations with the same amenities but with different admissions probabilities to schools. BDD mimics the ideal design by comparing locations that are within a narrow buffer around a school district boundary but on opposite sides. The identification assumption is that unobserved amenities are as good as random within a narrow buffer around a boundary. This assumption likely holds if other amenities are continuous in geography.[12]

We consider a narrow buffer that covers locations within 0.25 miles from a border at which a pair of school districts meet. Figure 1.5 illustrates this idea. Tables in subsection 1.9.2 present estimates with a 0.2-mile buffer. Table 1.8 presents summary statistics of student, housing, and Census block group characteristics of all sample in comparison to sample included in the BDD analysis. The differences in characteristics largely come from the fact that we exclude Staten Island since it consists of one school district. For example, Staten Island has larger number of White student, thus BDD sample has smaller share of White students (8.5%) than the full sample (12.5%).

The baseline regression is as follows.

[12]We do not apply BDD on attendance zone boundaries, because there is a concern about these boundaries' being determined by residents themselves. School district boundaries can be redrawn only every 10 years, and the decision is made at city level (New York State Law 2590-B). Meanwhile, attendance zone boundaries can be redrawn every year by the district council, whose members include parents and representative students. We still consider that admissions probability chances vary across attendance zones in the model estimation (Section 1.5).

$$y_i = \beta \underbrace{Q_{d(i)}}_{\text{district school quality}} + \theta_{b(i)} + f(r_i) + \epsilon_{id}\,. \tag{1.1}$$

The unit of observation i is a housing transaction record when y_i is the log house sales price. The unit of observation i is a Census block when y_i is the number and characteristics of middle-school-applying residents in the block. $b(i)$ is the boundary region fixed effect in which i is located. $f(r_i)$ is a local cubic control for distance to the boundary $b(i)$, which we allow to differ by whether the district in which i is included has higher school quality than the bordering district.

$Q_{d(i)}$ is district school quality, measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. In the model estimation, we consider multidimensional school "quality" measures and allow students to have heterogeneous preferences over measures. In this section, we use a one-dimensional measure for simplicity. Our choice of the mean test score is motivated by Abdulkadiroğlu et al. (2020)'s finding that parents of high school students in NYC do not value school effectiveness beyond the average test scores of students enrolled in a school.

The identification assumption is unlikely to hold if school district boundaries were drawn to divide already divided neighborhoods; even if they were exogenously drawn in the beginning, location amenities might have evolved differently over time on opposite sides of a boundary (Baum-Snow and Ferreira, 2015). We do two things to address these and to render causal interpretation of our estimates more plausible (Bayer et al., 2007; Kulka, 2019; Zheng, 2022).

First, we drop boundaries in which locations on opposite sides are likely to differ in access to amenities other than schools. Thus, we exclude boundaries aligned with a river, creek, park, highway, or borough boundary. Second, in subsection 1.9.2, we show that neither housing characteristics nor urban amenities change sharply at school district boundaries, which suggests that the identification assumption is plausible in our context.

Figure 1.3 presents estimates $\hat{\beta}$ for various outcomes. Tables in subsection 1.9.2 present coefficients plotted in Figure 1.3. For each outcome, we start from a simple

BDD specification (Equation 1.1). Then we present coefficients from specifications where we control for various covariates. In each panel, the coefficient from our preferred specification is in the rightmost.

District school quality increases the quality of schools to which residents are assigned The top left panel of Figure 1.3 reports $\hat{\beta}$ for the mean score of schools to which middle-school-applying residents in a Census block are assigned. A one standard-deviation increase in district school quality increases assigned schools' test scores of residents by 0.26 student-level standard-deviation (p-value < 0.01); we control for the resident's ethnicity, FRL status, and test score, to absorb differences in school applications and admissions probabilities explained by applicants' observable characteristics. This result implies that school district boundaries determine admissions probabilities to high-achieving schools, which establishes the first stage of the BDD.

District school quality increases housing prices The top right panel of Figure 1.3 reports $\hat{\beta}$ for the log sales price of a residential unit. Including this panel, we plot coefficients from specifications where we sequentially add housing characteristics, neighbor characteristics, and urban amenities for the rest of the panels. Given that housing characteristics and urban amenities do not change at a boundary (subsection 1.9.2), we control for those to increase the precision of our estimates. Meanwhile, we control for neighbor characteristics to account for the fact that households might have preferences over neighbors' ethnicity or median income. We interpret the estimate from a model with full controls to describe the effect of district school quality.

A one standard-deviation increase in district school quality increases housing sales price by 10% (p-value < 0.05). This implies that there is a higher demand for locations with higher admissions probability to better-performing schools.

We present $\hat{\beta}$ for house value and median gross rent from the ACS 5-year estimates in subsection 1.9.2. Estimates are 5.8% for both house value and median rent, although the estimate is only significant for median rent (p-value < 0.1).

While sold properties might not be representative of all properties, we prefer sales prices to these two alternatives because the ACS 5-year estimates are at Census block group level, which is too coarse to study a change of housing costs at boundaries. In subsection 1.9.2, we explain how we use the distribution of total population and houses across Census blocks within each block group to weigh Census block groups in obtaining $\hat{\beta}$.

District school quality attracts households with middle school applicants The bottom left panel of Figure 1.3 reports $\hat{\beta}$ for the number of middle school applicants residing in a Census block. A one standard-deviation increase in district school quality increases the number of middle-school-applying residents, with $\hat{\beta} = 0.79$ (p-value < 0.01). An average Census block has 3.5 middle-school-applying residents, and thus this is a 22% increase from the average. This result is robust to controlling for the total number of population in Census Block Group ($\hat{\beta} = 0.81$, p-value < 0.01) from the ACS 5-year estimate. Thus, we exclude an explanation that Census blocks with higher district school quality have a greater number of households with middle school applicants merely because those blocks have more houses. Estimates are presented in Table 1.11

District school quality attracts non-minority households more The two bottom panels of Figure 1.3 report $\hat{\beta}$ for the share of Black and Hispanic applicants among middle-school-applying residents in a Census block. A one standard-deviation increase in district school quality decreases the share of minority applicants by 6 percentage points (p-value < 0.01). An average Census block has 62% Black or Hispanic residents among middle-school-applying residents, and thus this is a 10% decrease from the average.

1.3.2 The Role of Commuting Distance in School Applications

Next, we show that while students apply to geographically proximate schools, the patterns are heterogeneous by students' characteristics and by the achievement level

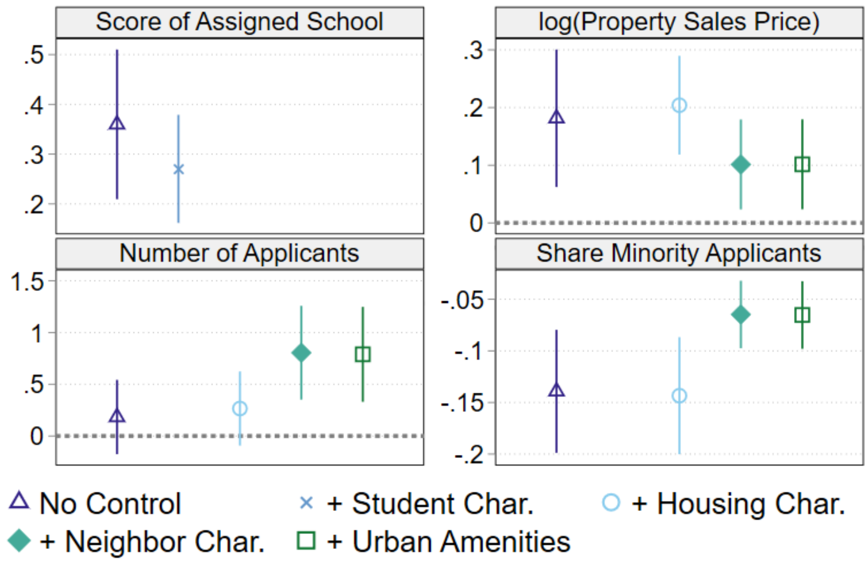


Figure 1.3: Estimated Effects of District School Quality on Residential Sorting

Note: The figure depicts the estimates (dots) and 95% confidence intervals (lines) of the coefficients of district school quality on various outcomes (β in Equation 1.1). The dependent variable in each panel is as follows (clockwise): (1) the mean score of the schools middle-school-applying residents in a Census block are assigned to, (2) the log sales price of a residential unit, (3) the number of middle-school-applying residents in a Census block, and (4) the share of Black and Hispanic applicants among those residents. In all panels, we plot the coefficient from a simple BDD specification (Equation 1.1) and coefficients from specifications that control for other variables. In the top left panel, we control for middle-school-applying residents' ethnicity, FRL status, and test score. In the rest of the panels, we sequentially add housing characteristics, neighborhood characteristics, and urban amenities. Standard errors are clustered at school district level. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

of schools near their residential locations. We run the following linear probability model:

$$100 * \mathbb{1}(\text{Top3})_{ij} = \alpha d_{\ell_i j} + d_{\ell_i j} Z_i \beta + \delta \mathcal{J}_i + \epsilon_{ij} \,. \tag{1.2}$$

$\mathbb{1}(\text{Top3})_{ij}$ is an indicator for whether student i lists school j in her top three choices. j is a school for which student i is eligible. We multiply $\mathbb{1}(\text{Top3})_{ij}$ by 100 to interpret coefficients as percentage point changes. $d_{\ell_i j}$ is the driving distance in miles between school j and student i's residential census block ℓ_i. Z_i is a vector of student characteristics. α represents the association between distance to a school and the propensity of students to list the school as their top choice. β shows how that association changes by students' characteristics. To account for the fact that the probability of choosing a specific school as the top choice mechanically decreases when the number of eligible options increases, we control for the total number of schools for which i is eligible ($\mathcal{J}_i$). We cluster standard errors at student level.

Columns (1) and (2) in Table 1.1 demonstrate that students are 2.4 percentage points less likely to rank a school that is 1 mile farther away as their top 3 choices (p-value < 0.01). Minority students seem to be less responsive to distance in column (1) ($\beta = 0.262$, p-value < 0.01). In column (2), we further control for the mean test score of the three closest schools from student i's residential Census block. Students are even less likely to apply to schools farther away when nearby schools have higher quality (column (2), $\beta = -0.286$, p-value < 0.01). Importantly, controlling for the quality of nearby schools reduces the coefficient of the minority dummy by two-thirds ($\beta = 0.082$, p-value < 0.01). This pattern, whereby students from disadvantaged location travel farther to schools that are higher performing than schools in their residential location, coincides with what has been reported in previous studies (Burdick-Will, 2017; Corcoran, 2018).

Motivated by these patterns, we model households as considering not only commuting distances but also other school characteristics. We also allow households to have heterogeneous commuting costs.[13]

[13]This pattern is not explained by the difference in the number of schools proximate to their

	(1)	(2)
$d_{\ell_i j}$	-2.460	-2.411
	(0.013)	(0.012)
$d_{\ell_i j} \times \mathbb{1}(\text{Minority})_i$	0.262	0.083
	(0.011)	(0.011)
$d_{\ell_i j}$ × Quality of the three closest schools$_i$		-0.286
		(0.007)
N	1,745,513	1,745,513
R2	0.062	0.063
Dep. var mean	7.895	

Note: The dependent variable is a dummy if student i listed school j as one of their top three choices, multiplied by 100 for ease of interpretation. Pairs of a student and an eligible school within 10 miles from the student's residential Census block are included. The fastest driving distance between a school and a Census block is calculated using Open Route Services. A student is a minority if she is Black or Hispanic. Column (2) controls for the mean test score of the three closest schools from i's residential Census block ℓ_i. All columns control for the total number of schools a student is eligible for and the interaction of distance and student's standardized test scores. Standard errors in parentheses are clustered at the student level.

Table 1.1: Commuting Distances and the Propensity of Listing as Top 3

1.4 A Model of Location Choice, School Choice, and Enrollment Decision

We model households' sequential decisions of residential locations, school applications, and enrollment decisions. Location choices affect school applications and assignments through two channels. First, distances to schools vary by residential location, which affects students' school applications. Second, applicants are ranked based on location-based priority rules, and thus two students from different locations who are otherwise similar face different admissions probabilities. Households take these two channels into account when choosing which location to reside in.

The model is guided by two key parameters. The first is access-to-school prefer-

residential location. Students whose proximate schools are lower achieving have more schools proximate to their residential location. We present a histogram showing this result in subsection 1.9.2.

ence α^u. It is the weight households put on access-to-school utility that captures both commuting distances and admissions probability, relative to other location amenities. It also governs the extent to which counterfactual location-based priority rules would induce households to resort across locations. The second is commuting cost β^d, which affects to what extent students apply to schools that are farther away given their location choices as opposed to applying to only nearby schools. Together with α^u, it shapes the spatial distribution of households; for example, with infinite commuting costs and strictly positive access-to-school utility, households would choose locations closer to the schools they would apply to.

Next, we discuss our model in greater detail.

1.4.1 Household Preference, School Assignment, and Timeline

Household Heterogeneity and Preferences We use “household, applicant," and “student” interchangeably and model the unitary decision of a household. Household i is heterogeneous in both observable and unobservable (to the researcher) characteristics, denoted as Z_i and γ_i, respectively. Observable characteristics Z_i include students' race/ethnicity, poverty status (proxied by free and reduced-price lunch eligibility), and test score prior to their middle school application. Unobserved (discrete) type γ_i (Heckman and Singer, 1984) captures the fact that school characteristics may be valued differently by observably similar households.

i's utility from living in location ℓ and attending school j is

$$\underbrace{V_i(\ell;\eta_{i\ell})}_{\text{utility from location}} + \alpha^u \underbrace{U_i(j,\ell;\varepsilon_{ij})}_{\text{utility from school} \times \text{location}} . \tag{1.3}$$

We parameterize each component as follows:

$$V_i(\ell;\eta_{i\ell}) = \underbrace{W_\ell'}_{\text{location char.}} \alpha_i^W + \underbrace{p_\ell}_{\text{housing cost}} \alpha_i^p + \underbrace{\xi_\ell}_{\text{unobserved amenities}} + \underbrace{\eta_{i\ell}}_{\text{i.i.d. EVT1}}, \tag{1.4}$$
$$\text{where } \alpha_i^k = \alpha^{k0} + \underbrace{Z_i'}_{\text{student char.}} \alpha^{kz}, \text{ for } k = p, W.$$

$$U_i(j,\ell;\varepsilon_{ij}) = \underbrace{X_j'}_{\text{school char.}} \beta_i^X + \underbrace{d_{\ell j}\beta_i^d}_{\text{commuting cost}} + \underbrace{\varepsilon_{ij}}_{\text{i.i.d. EVT1}}, \tag{1.5}$$
$$\text{where } \beta_i^k = \beta^{k0} + \underbrace{Z_i'}_{\text{student char.}} \beta^{kz} + \underbrace{\gamma_i^k}_{\text{unobserved type}}, \text{ for } k = d, X.$$

W_ℓ is a vector of location observable characteristics, p_ℓ is the housing cost, Z_i is the vector of student observable characteristics, X_j is the vector of school characteristics, and $d_{\ell j}$ is the fastest driving distance between location ℓ and school j.

In addition, ξ_ℓ represents unobservable location amenities that are shared across households. $\gamma_i = (\gamma_i^X, \gamma_i^d)$ is the vector of student i's unobserved tastes over school characteristics and distance to schools, and $\eta_{i\ell}$, ε_{ij} are idiosyncratic preferences shocks over locations and schools. $\eta_{i\ell}$ and ε_{ij} are mutually independent and follow i.i.d extreme value type 1 distribution.

i's utility from living in location ℓ and attending an outside option ϑ is

$$\underbrace{V_i(\ell;\eta_{i\ell})}_{\text{utility from location}} + \underbrace{U_i^\vartheta(\vartheta;\varepsilon_{i\vartheta})}_{\text{utility from outside option}}. \tag{1.6}$$

We consider two outside options, non-public schools ϑ^{np} and public charter schools ϑ^c.[14] Non-public schools ϑ^{np} includes private schools, homeschooling, or moving out of NYC.[15] We further allow students to have heterogeneous preferences

[14]Public charter schools are not parts of the centralized school choice system and they have separate admissions processes.

[15]Although we observe that a student is not enrolled in a public school in NYC, we do not know which non-public option a student chooses.

for outside options based on their observable characteristics. For example, non-minority students might assign higher value to non-public schools than their peers. Mathematically,

$$U_i(\vartheta;\varepsilon_{i\vartheta}) = \beta_i^{\vartheta} + \underbrace{\varepsilon_{i\vartheta}}_{\text{i.i.d. EVT1}} \tag{1.7}$$
$$\beta_i^{\vartheta} = \beta_0^{\vartheta} + \underbrace{Z_i}_{\text{student char.}} \beta_z^{\vartheta}, \text{ where } \vartheta = \vartheta^c, \vartheta^{np}.$$

$\varepsilon_{i\vartheta}$ follows an i.i.d extreme value type 1 distribution.

School Assignment Next, we briefly discuss how schools rank applicants. As discussed in Section 1.2, priority groups are largely determined by students' residential location ℓ. The tie-breaker within priority groups is either a lottery or a school-specific aggregation of students' pre-middle-school academic measures. We capture programs' ranking over students with a priority score, $c_{ij}(\ell)$. This is the sum of an integer $g_{ij}(\ell)$ that corresponds to priority groups and decimal point $\tau_{ij} \in [0, 1]$ that corresponds to tie-breakers.[16] Tie-breakers are either school-specific aggregation of students' academic measures or random lottery numbers ρ. The higher a student's $c_{ij}(\ell)$, the higher her admissions chance.

How priority groups are determined is public information. When it comes to school-specific aggregation of students' academic measures, we know which inputs a school uses—such as GPA, statewide test score, and punctuality—and the aggregated scores among its applicants. However, the exact function that schools use to construct these measures are unknown. We estimate school-specific linear functions of measures using a latent model and assume households form expectations $(\hat{c}_{ij}(\ell))$ in the same way; details are in subsection 1.9.3. Given students' rank-ordered list and priority scores, the city assigns students to at most one program using the SPDA algorithm. See subsection 1.9.1 for a detailed explanation of the SPDA procedure.

[16] For a program j with three priority groups, students in the first priority group have $g_{ij} = 3$. The second and the third priority groups' students have $g_{ij} = 2$ and 1, respectively. If a student is ineligible for program j, $g_{ij} = -\infty$

Cutoff for each school $\bar{c}_j$ is given as the $\min\{c_{ij} : i \in \mathcal{I}_j\}$, where $\mathcal{I}_j$ is a set of students admitted to program j if the capacity of j is filled, and $-\infty$ otherwise. We assume the market is large enough (70,000 students) that an individual student considers the cutoffs as given (Fack et al., 2019a; Agarwal and Somaini, 2020; Calsamiglia et al., 2020).

Timing Figure 1.4 summarizes the timeline of the model and households' information at each stage. Households make choices on blue dots. They have full information on their own observable characteristics (Z_i) and those of schools (X_j) and residential locations (W_ℓ, p_ℓ), as well as locations' unobserved amenities (ξ_ℓ), throughout all stages.

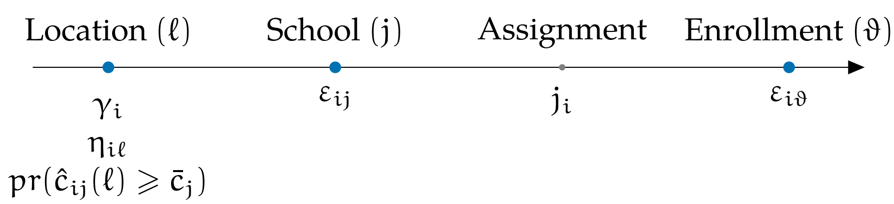


Figure 1.4: Timeline and Information Set

Note: Households make choices on blue dots. They have full information on their own observable characteristics (Z_i) and those of schools (X_j) and residential location (W_ℓ, p_ℓ), as well as the shared neighborhood unobserved amenities (ξ_ℓ) throughout all stages. γ_i is the vector of student i's unobserved tastes over school characteristics. $\eta_{i\ell}$, ε_{ij}, and $\varepsilon_{i\vartheta}$ are idiosyncratic preferences shocks over locations, schools, and outside options respectively. j_i is the assignment result. $pr(\hat{c}_{ij}(\ell) \geqslant \bar{c}_j)$ is the predicted admissions probability.

Households know their unobserved tastes over school characteristics (γ_i) from the beginning of the location choice stage, so these unobserved preferences influence their residential choice. This becomes a source of bias in estimating commuting costs if we estimate school preference without correcting for the selection into locations. For example, a household that values school safety more than other observably similar households would choose locations that assure higher location-based admissions probability for safer schools. This household would apply to only nearby schools because it already lives near its safer schools, but a model that does not correct this selection would mistakenly justify such behavior with a high commuting cost. Together with unobserved taste, households observe their idiosyncratic preference shocks over locations $(\eta_{i\ell})$ and form predictions on admissions probabilities to schools, $\text{pr}(\hat{c}_{ij}(\ell) \geqslant \bar{c}_j)$.[17]

At the beginning of the school choice stage, student i observes her preference shock over programs, ε_{ij}. To sum up, households know know unobserved tastes γ_i but not idiosyncratic shock ε_{ij} when deciding where to live. Once the assignment is realized, they know the exact assignment result j_i. The preference shock over outside options $\varepsilon_{i\vartheta}$ is realized at the enrollment-decision stage to rationalize the fact that 7.74% of students assigned to their top choice enroll in outside options. $\varepsilon_{i\vartheta}$ is either an income shock that affects households' affordability for private schools (Calsamiglia et al., 2020) or charter school lotteries realized after the application stage is complete. The idiosyncratic shock ε_{ij} over assigned school does not change in the enrollment-decision stage.[18]

1.4.2 Household's Problem

Next, we describe the household's problem corresponding to the blue dots in Figure 1.4, which we solve backward.

Stage 4: Enrollment Residential locations and assignment results are set in previous stages. Given those, students decide whether to enroll in their assigned school, or the non-public option, or a public charter school to maximize their utility:

[17]The admissions cutoffs $\bar{c}_j$ households use at this stage are calculated using observed school application, which is a function of students' preference shocks over programs ε_{ij} that are realized in the next period. The large market assumption establishes the internal consistency—i.e., the admissions cutoffs are determined in the large market, and $\bar{c}_j$ are consistent estimators of those.

[18]An alternative model is such that students draw new shocks on assigned schools at the enrollment stage and the final shock is a weighted sum of the old and the new shock. However, it is impossible to tell to what extent the idiosyncratic shock ε_{ij} is time-invariant, since all other choices from the application stage are forgone except for j_i. That is, this alternative model would generate the same school application list and enrollment decision. Such model would have been possible if students had more than two options that are relevant at both the application stage and enrollment stage.

$$U_i^*(\ell_i) \equiv \max\left\{U_i(j_i(\ell_i), \ell_i; \varepsilon_{ij}), U_i^{\vartheta}(\vartheta^{np}; \varepsilon_{i\vartheta}), U_i^{\vartheta}(\vartheta^{c}; \varepsilon_{i\vartheta})\right\}, \tag{1.8}$$

where j_i is the assignment outcome from the assignment stage and ℓ_i is the location chosen in the previous stage.

Stage 3: Assignment Students are passive as their assigned school j_i is determined by their priority score at each program and admissions cutoffs, given their ROLs from the previous stage. Mathematically,

$$j_i(\ell_i) \equiv f(\underbrace{\mathcal{ROL}_i(\ell_i)}_{\text{application list}}, \underbrace{c_{ij}(\ell_i; \rho)}_{\text{priority score}}, \underbrace{\bar{c}_j}_{\text{cutoff vector}}). \tag{1.9}$$

Stage 2: Application We assume that students submit an ROL following their true preference order up to their fallback options. The fallback option is the school a student is assigned to when rejected by all programs on her ROL—either predesignated zoned school or an undersubscribed school in her school district. The middle school choice system in NYC uses the Deferred Acceptance algorithm, in which students can list as many schools as they want, which jointly renders truth-telling—ranking schools based on one's true preference order—a weakly dominant strategy (Gale and Shapley, 1962).

Stage 1: Residential Location Choice Given the solution in the subsequent period, household i chooses the location that solves

$$\max_{\ell} \underbrace{V_i(\ell; \eta_{i\ell})}_{\text{utility from location}} + \underbrace{\alpha^u \, \mathbb{E}_{\varepsilon_{ij}, \rho, \varepsilon_{i\vartheta}} U_i^*(\ell)}_{\text{expected utility from enrolled school given location}}, \tag{1.10}$$

where $U_i^*(\ell)$ is the utility from enrolled school (Equation 1.8). This is is location dependent because locations decide commuting costs and admissions probabilities, and as a result, which school student i enrolls in. Households form an expectation over $U_i^*(\ell)$, since they do not know their idiosyncratic preference shocks over schools and outside options (ε_{ij} and $\varepsilon_{i\vartheta}$) as well as their lottery number ρ_i.

1.4.3 Equilibrium

To define the school assignment equilibrium, we extend the supply and demand characterization of Azevedo and Leshno (2016a).

Definition 1.1. *An equilibrium is a pair of decisions* $\{\ell_i, \mathcal{ROL}_i\}$ *for each* i *and a vector of admissions cutoffs* $\{\bar{c}_j\}_{j=1}^J$ *where*

1. *Given cutoffs* $\{\bar{c}_j\}_{j=1}^J$ *and the first-stage choice* ℓ_i, $\{\mathcal{ROL}_i\}$ *is the school application list based on* i*'s true preference order up to their fallback option.*

2. *Given* $\{\bar{c}_j\}_{j=1}^J$, ℓ_i *solves* i*'s problem Equation 1.10 for each* i.

3. *Admissions cutoffs clear the market; i.e.,* $S_j \geqslant D_j(\{\bar{c}_{j'}\}_{j'=1}^J)$ *for each* $j \in J$. S_j *is capacity of school* j *and* $D_j(\{\bar{c}_{j'}\}_{j'=1}^J)$ *is the aggregate demand for school* j *given the cutoffs* $\{\bar{c}_{j'}\}_{j'=1}^J$.

Aggregate demand for schools can be further simplified by using the fact that when students are truth-telling, the realized matching is stable—i.e., each student is matched to her favorite feasible school. Details are in subsection 1.9.3.

1.4.4 Discussion

Truth-telling We consider truth-telling a reasonable assumption in our context. There are well-known factors that make this assumption less plausible: (1) list-length restriction (2) limited consideration set (3) application cost. First, there is no list-length restriction in our setting. In a setting with this restriction (Luflade, 2018; Son, 2020), truth-telling is no longer a weakly dominant strategy when students really want to be assigned to some school (Haeringer and Klijn, 2009a).[19] Second, students are given a customized list of all eligible schools with an average of about 30 schools. This stands in contrast to settings in which they have to construct a consideration set out of hundreds of options, where they are unlikely to consider

[19] List-length restrictions were introduced in NYC's middle school choice in years more recent than our setting.

all options in their choice set when deciding which school to apply to (Ajayi and Sidibe, 2020; Son, 2020). Third, both monetary and psychological application cost are relatively low. There is no application fee. Also, they can add one more school to their application list just by marking the ranking to the customized list that they received.

Even though assuming truth-telling is reasonable in our context, it is still a weakly dominant strategy (Artemov et al., 2017; Che et al., 2022). For example, "skipping the impossible" yields the same assignment results, and detecting impossible options is feasible given that each school's capacity and the number of previous year's applicants are public information. Instead of imposing truth-telling assumption, one can estimate the model based on stability (Fack et al., 2019a; Agarwal and Somaini, 2020; Hahm and Park, 2022), which rely on assignment results rather than ranking strategies, in subsection 1.9.4. While imposing a weaker assumption, this estimation strategy loses the precision of estimates by focusing only on the assignment outcome instead of the full list.

Asymmetry in Utility from Location and School In our model, utility from locations includes unobserved amenities shared by households but no household-specific unobserved tastes. Meanwhile, utility from schools includes household-specific unobserved taste but no unobserved quality shared by households (Equation 1.4).

These modeling choices are largely driven by the motivation to obtain unbiased estimates of two key parameters—access-to-school preference α^u and commuting cost β^d—while keeping the estimation tractable. Access-to-school preference α^u will be biased if unobserved location amenities are correlated with access-to-school utility.[20] Meanwhile, households' sorting into locations based on household-specific

[20]Moreover, we lack variation to identify household-specific unobserved tastes. For example, Bayer et al. (2016) sets up a dynamic location choice model and uses the panel structure of the data to identify households' unobserved attachment to a specific location. Or, Barwick et al. (2021) constructs household-specific location choice set by leveraging that they observe when each household bought the house.

unobserved school tastes biases commuting cost β^d.[21]

Utility from the Outside Option Utility from the outside option is not a function of school characteristics because of a lack of data on schools outside the system, especially non-public options. It can also be a function of location. By abstracting away from it, our location demand estimates might capture the unequal geographic distribution of outside schooling options. For example, locations with higher median household income would have more private schools nearby, and the estimated preference over neighbors' income in Section 1.5 might capture households' preference over geographic proximity to non-public options. We assume the geographical distribution of outside options does not change under the counterfactual scenario.[22]

1.5 Estimation Procedure and Results

1.5.1 Identification of Key Parameters

We discuss the identification of the two key parameters of the model: access-to-school preference α^u and commuting cost β^d. We also discuss the identification of price coefficient α^p.

The biggest concern regarding credibly identifying access-to-school preference α^u is to distinguish it from preferences on unobserved location amenities (ξ_ℓ). To this end, we use variation from our boundary discontinuity design. Similar to Section 1.3, the identification assumption is that the unobserved amenities are as good as random within a narrow buffer around a boundary. Meanwhile, the access-to-school utility sharply changes at a boundary, since 70% of schools give eligibility or higher priority to students from the same school district (Section 1.2), and there is marked heterogeneity in school characteristics across districts. So

[21]For example, Allende (2019) and Abdulkadiroğlu et al. (2020) estimate unobserved school quality to obtain causal estimates on how much households value peer quality aside from other school factors such as the building quality.

[22]See, for example, Dinerstein and Smith (2021) to see how private schools' entry and exit decisions can be affected by public school policies.

intuitively, seeing that households are more likely to live in the side of a district boundary with higher admissions probabilities to schools whose characteristics are more desirable (Section 1.3) would lead to a larger value of α^u.

To obtain an unbiased estimate of commuting cost β_i^d, we need to account for the fact that households choose locations based on their unobserved school demand γ_i. When students apply to nearby schools, we need to identify to what extent this is explained by commuting costs as opposed to households' residential sorting in order to be assigned higher priority by the schools they prefer. If residential sorting arises only from households' observed characteristics, we can obtain unbiased commuting costs by controlling for those characteristics in estimating school preference without fully modeling residential sorting. Thus, previous papers have assumed that idiosyncratic preference shocks and unobserved tastes over schools are independent of distances to school conditional on student observable characteristics—i.e., $(\varepsilon_{ij}, \gamma_i) \perp d_{l_i j}|Z_i$ (e.g., Agarwal and Somaini, 2018a; Laverde, 2020). By modeling and jointly estimating location and school choice, we relax this assumption and allow an individual's unobserved type γ_i to be correlated with distances to schools—i.e., $\varepsilon_{ij} \perp d_{l_i j}|Z_i$.

Moreover, we need to identify unobserved type γ_i to correct for households' selection into locations based on it. Whereas the different applications of two observably identical students can be explained by either unobserved tastes (γ_i) or idiosyncratic preference shock (ε_{ij}), these two components can be disentangled for two reasons. First, unobserved tastes are student-specific but the preference shock is independent across schools within each student. We observe students' full application lists. To what extent characteristics among the schools on a student's ROL are correlated helps to identify the unobserved type separate from the idiosyncratic shock (Bhat, 2000; Berry et al., 2004). Second, while households choose residential locations knowing their unobserved taste, the idiosyncratic shock is realized after location choice. Among observably similar students, variation across residents of different locations pins down unobserved taste, while variation among residents from the same location are captured by the idiosyncratic preference shock.[23]

[23]In principle, we can even divide unobserved taste into two components: individual-specific

The final parameter to identify is the price coefficient, α_p, since housing cost (p_ℓ) is likely to be correlated with location unobserved amenities (ξ_ℓ). We instrument for housing cost of a location with the land use of other locations that are (1) 2 miles away from the location (2) but within 3 miles (Bayer et al., 2007; Barwick et al., 2021; Davis et al., 2021). Given a location, other locations that are far away are unlikely to share its unobserved amenities (exclusion restriction). However, the land use of other locations that are near enough to the location could affect its housing cost if people decide where to live among those locations (relevance restriction).

1.5.2 Estimation Procedure

Challenge 1: Granularity of Location With over 38,000 Census blocks in NYC, estimating the model at block level might decrease the precision of estimates by having too many parameters relative to the data (Dingel and Tintelnot, 2020).[24] But we still aim to estimate the mean utility of location ($\delta_\ell = W_\ell \alpha^{W0} + p_\ell \alpha^{p0} + \xi_\ell$) to account for the endogeneity of access-to-school and housing price (Berry et al., 1995). To this end, we define neighborhoods—a unit of residential location—by merging Census blocks.

Two Census blocks are in the same neighborhood if they satisfy the following criteria. First, they are in the same cluster when we group Census blocks based on the distance to all schools using k-mean clustering, with k of 1,000. Second, they share the same location-based admissions probability to all schools. Third, they are either both within the 0.2-mile buffer of a school district boundary or neither is.

With this procedure, we aggregate 38,798 Census blocks into 2,778 neighborhoods. In comparison, there are 2,165 Census tracts in NYC. Figure 1.5 shows the map of neighborhoods defined by this procedure. The darker shaded neighbor-

and realized at the location choice stage versus individual-specific and realized at the application stage.

[24]There are also many Census blocks with no middle-school-applying residents or housing transaction records during the time of the study.

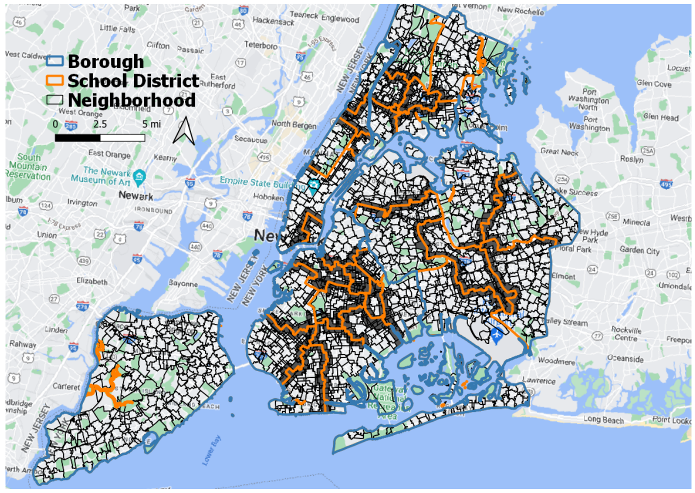


Figure 1.5: Defined Neighborhoods

Note: We aggregate 38,798 Census blocks into 2,778 neighborhoods using the procedure described in Subsection 1.5.2. The darker shaded neighborhoods along school district boundaries (in orange) are those to which we apply BDD to identify access-to-school preference α^u.

hoods along school district boundaries (in orange) are those to which we apply BDD to identify access-to-school preference α^u.

Challenge 2: Computational Burden from Joint Estimation We aim to jointly estimate all stages of the model to address the selection into locations. Full information maximum likelihood (FIML) involves calculating a large Hessian matrix (Train, 2009), which renders the computation infeasible. See subsection 1.9.4 for details on FIML.

To circumvent the computational burden, we employ the expectation-maximization algorithm with a sequential maximization step (ESM) proposed by Arcidiacono and Jones (2003).[25] In summary, the idea is to (1) reformulate the full information likelihood function into additive separable terms, each of which represents the likelihood of each stage; (2) update estimates of each stage; and (3) iterate the

[25]Dempster et al. (1977) and Train (2009) show that solving the EM algorithm is identical to solving maximum likelihood estimation (MLE), and Arcidiacono and Jones (2003) prove the consistency of estimates with a multi-stage model.

procedure until convergence.

The expectation function (reformulation) for the household i is a sum of the log of the likelihood for each stage weighted by the conditional probability of its being each unobserved type, given the school application and location choice observed in the data. Then we take the sum across i's expectation function.

$$\begin{aligned}
\mathcal{E}(p,\gamma,\theta|\hat{q},\hat{\gamma},\hat{\theta}) &= \Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) \log q_k && (1.11)\\
&+ \Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) \log P^{LC}(x_i;\theta^{EC},\theta^{SC},\theta^{LC},\gamma_k) && (1.12)\\
&+ \Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) \log P^{SC}(x_i;\theta^{SC},\gamma_k) && (1.13)\\
&+ \Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) \log P^{EC}(x_i;\theta^{EC},\theta^{SC},\gamma_k)\,. && (1.14)
\end{aligned}$$

$q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta})$ is the conditional probability of being type k given data x_i, calculated using Bayes' rule. θ^{LC}, θ^{SC}, θ^{EC} are the set of location, school, and outside option preference parameters, respectively. γ_i is the unobserved taste, and q_k is the unconditional probability of each type k. $\theta = \{(\theta^{LC},\theta^{SC},\theta^{EC}),\{\gamma_k, q_k\}_k\}$ is the full set of parameters to be estimated. P^{LC}, P^{SC}, P^{EC} are the likelihood of location choice, school choice, and enrollment choice, respectively. Likelihood functions are presented in subsection 1.9.4.

Then we update the guess on each element of θ sequentially by maximizing each line of the expectation function. Starting from an initial guess, we iterate the updating process until the guess of θ converges. We used squared extrapolation methods (see Varadhan and Roland, 2008) to make convergence faster. See subsection 1.9.4 for the cookbook of the iteration process.

We update θ^{SC}, θ^{EC}, and γ using maximum likelihood estimation to obtain efficient estimates of γ by exploiting full information in application lists.

Meanwhile, we update θ^{LC} using method of moments estimation to deal with the endogeneity of price and access-to-school utility (Berry et al., 1995). Location preference parameters include those that govern heterogeneous preferences $(\alpha^{Wz}, \alpha^{pz})$; common preferences $(\alpha^{W0}, \alpha^{p0})$; and the access-to-school preference α^{u}. To estimate $(\alpha^{Wz}, \alpha^{pz})$, we match the first-order condition of location choice likelihood P^{LC} (presented in subsection 1.9.4) with respect to α^{Wz} and α^{pz},

$$\underbrace{\Sigma_i W_{\ell_i} z_i^r}_{\text{cov. of } W \text{ and } z \text{ in the data}} = \underbrace{\Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) \Sigma_\ell P^{LC}(\ell;\hat{\theta}^{SC},\hat{\theta}^{EC},\theta^{LC},\hat{\gamma}_k) W_\ell z_i^r}_{\text{predicted cov. of } W \text{ and } z}, \quad (1.15)$$

$$\underbrace{\Sigma_i p_{\ell_i} z_i^r}_{\text{cov. of } p \text{ and } z \text{ in the data}} = \underbrace{\Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) \Sigma_\ell P^{LC}(\ell;\hat{\theta}^{SC},\hat{\theta}^{EC},\theta^{LC},\hat{\gamma}_k) p_\ell z_i^r}_{\text{predicted cov. of } p \text{ and } z}, \quad (1.16)$$

where z^r is each element of observed household characteristics Z—e.g., the minority dummy.

To obtain the remaining parameters $(\alpha^{W0}, \alpha^{p0})$, α^u, we first search the mean utility of location $\delta_\ell = W_\ell \alpha^{W0} + p_\ell \alpha^{p0} + \xi_\ell$ that satisfies the first-order condition of P^{LC} with respect to δ_ℓ, [26]

$$\underbrace{\Sigma_i \mathbb{1}(\ell_i = \ell)}_{\text{observed share}} = \underbrace{\Sigma_i \Sigma_k q(k|x_i;\hat{q},\hat{\gamma},\hat{\theta}) P^{LC}(\ell;\hat{\theta}^{SC},\hat{\theta}^{EC},\theta^{LC},\hat{\gamma}_k)}_{\text{predicted share}}, \forall \ell . \quad (1.17)$$

Finally, we get $(\alpha^{W0}, \alpha^{p0}, \alpha^u)$ by targeting the following conditions"

$$E(\xi_\ell \hat{p}_\ell^{IV}) = 0 \quad \text{(Price IV)}$$

$$E(\xi_\ell \mathbb{1}(\text{right side of BD})_\ell | BD_\ell, \mathbb{1}(\ell \in \mathcal{B}(BD_\ell; 0.25\text{mi.})) = 0 . \quad \text{(BDD IV)}$$

where $\hat{p}_\ell^{IV}$ is a vector of other observed location characteristics W_ℓ and price IV. $\mathcal{B}(BD_\ell; 0.25)$ is the buffer around each boundary BD with a radius of 0.25 mile. The procedure consists of the outer loop that searches parameters that satisfy Equation 1.15, 1.16, Price IV, and BDD IV and the inner loop that searches δ_ℓ that satisfy Equation 1.17. We present price IV regression results in subsection 1.9.4.

[26]The process is accelerated by Newton's nonlinear root-finding algorithm. We thank Jean-François Houde for sharing his code.

1.5.3 Estimation Results

Demand Estimates Estimates in Table 1.2 have expected signs. Households prefer locations with higher access-to-school utility (EU, 1.419) and lower housing costs. They prefer schools that are higher achieving and safer. There is homophily (i.e., preference for one's same race and FRL status) in both location and school preferences.

Willingness to pay To better interpret estimates, we calculate households' willingness to pay (WTP) for school and location characteristics in Table 1.3. WTP for a one-unit increase in location characteristics for household i is given by $\frac{\alpha_i^W}{\alpha_i^P}$. WTP for a one-unit increase in school characteristics of *all* schools in a school district is the sum of $\frac{\alpha^u}{\alpha_{p_i}} \frac{\partial EU_{i\ell}}{\partial X_j}$ across school js in a school district of location ℓ. This is household WTP to ensure the increase in characteristics of the assigned school[27] and is a function of location; thus we take the average across locations. We further convert WTPs in monetary terms by multiplying them by \$1,366, the mean of median gross monthly rent at the Census tract from 2014 5-year ACS estimates.

For some characteristics, households uniformly agree on what makes a location or a school more desirable. Both the 25th and 75th percentiles of households are willing to pay a positive amount for an increase in the median income of neighbors, mean test score of schools, and safety of schools.[28] For other characteristics, there is marked heterogeneity in preferences. For an increase in the minority share among neighbors or school peers, some households are willing to pay a positive amount while others must be compensated to stay indifferent.[29]

[27] $\frac{\partial EU_{i\ell}}{\partial X_j}$ can be simplified to $\beta_i^X \mathtt{Prob}_{ij}(\ell)\mathtt{Prob}(j \in \mathcal{J}_i(\ell;\rho_i))$ where $\mathtt{Prob}_{ij}(\ell)$ is the probability of j's being the most preferred feasible option for student i when she lives in ℓ. $\mathtt{Prob}(j \in \mathcal{J}_i(\ell;\rho_i))$ is the probability of j being i's feasible choice when she lives in ℓ.

[28] WTP for a one-standard-deviation increase in schools' test score is 11.3%. Ours is slightly higher than the range reported by previous papers (3%-10%) that study households' WTP for a test score increase in *one* school such as a zoned school or a charter school (Black, 1999; Bayer et al., 2007; Zheng, 2022). In contrast, we consider a test score increase for *all* schools in a district.

[29] Bayer et al. (2016) estimate that for a 10-percentage-point increase in the fraction of White neighbors, an average White family in the San Francisco Bay Area is willing to pay \$2,428 annually in 2000 dollars from their dynamic location choice model, and \$1,901 from their static model. Our

	Main	Additional Effects				
	Type1	Type2	Type3	Black/Hisp	FRL	Low-achieving
Panel A: Neighborhood Demand						
log(SalesPrice)	-2.039	-	-	-0.015	-0.048	-0.035
	(1.383)			(0.283)	(0.360)	(0.272)
Frac. Black or Hisp.	-3.459	-	-	3.933	-0.119	0.208
	(2.025)			(1.727)	(1.671)	(1.672)
log(Med. HH Income)	2.848	-	-	-0.161	-1.201	-0.303
	(1.991)			(1.545)	(1.891)	(2.141)
Med. Time to Work (hr)	17.676	-	-	-27.016	13.242	15.394
	(359.156)			(278.872)	(402.963)	(474.360)
Med. Time to Work^2 (hr)	-13.456	-	-	17.269	-8.257	-10.207
	(247.912)			(191.749)	(278.871)	(323.379)
EU	1.419	-	-	-	-	-
	(1.276)	-	-	-	-	-
Panel B: School Demand						
Mean test score	0.121	0.256	0.187	0.134	-	-0.253
	(0.052)	(0.074)	(0.469)	(0.029)	-	(0.028)
Frac. Black or Hisp.	-1.722	-0.501	0.159	1.958	0.117	0.216
	(0.612)	(1.118)	(5.442)	(0.417)	(0.281)	(0.223)
Frac. FRL	-0.771	-0.356	1.020	-0.540	0.882	-0.162
	(0.865)	(1.041)	(5.097)	(0.429)	(0.220)	(0.213)
Non-safety	-0.059	0.003	0.027	0.018	-	-
	(0.008)	(0.012)	(0.057)	(0.010)	-	-
Commuting Cost (mi.)	0.221	1.085	9.136	-0.084	-0.010	-0.046
	(0.032)	(0.054)	(1.757)	(0.045)	(0.021)	(0.014)
Prob.	0.352	0.631	0.017	-	-	-
	(0.127)	(0.083)	(0.055)	-	-	-
Panel C: Outside Option						
Non-public	-1.225	-	-	-0.136	-0.800	-0.200
	(0.174)			(0.135)	(0.125)	(0.134)
Public Charter	-3.767	-	-	1.883	0.173	-0.059
	(0.256)			(0.207)	(0.152)	(0.124)

Note: Standard errors in parentheses are calculated from 75 bootstrapped samples. Columns are for students' heterogeneity and rows are for school and neighborhood characteristics. FRL stands for free or reduced-price lunch eligibility. The fastest driving distance to a school is calculated using the Open Route Service. School non-safety measure is constructed by running a principal component analysis on crime incidence of different categories at each school building. See subsection 1.9.4 for details.

Table 1.2: Demand Estimates

	Std of Var.	WTP		
		p25	median	p75
Panel A: Neighborhood Characteristics				
Frac. Black or Hisp.	0.341	-704	32	71
log(Med. HH Income)	0.429	303	379	595
Med. Time to Work (min, daily)	11.52	-236	-181	-114
Panel B: School Characteristics				
Time to School (min, daily)	51	-752	-384	-171
Mean test score	1.002	89	154	238
Frac. Black or Hisp.	0.28	-355	-25	11
Frac. FRL	0.173	-98	-83	-53
Safety	4.228	94	121	149

Note: The unit of willingness to pay is the mean of median gross monthly rent at the Census tract from 2014 5-year ACS estimates, $1,366. We use the standard deviation of distance to the assigned schools across students. For other school characteristics, we calculate the standard deviation across schools. FRL stands for free or reduced lunch eligibility. The fastest driving time to school is calculated using the Open Route Service. School safety is constructed by running a principal component analysis (PCA) on crime incidence of different categories at each school building. subsection 1.9.4 has more details on the PCA result.

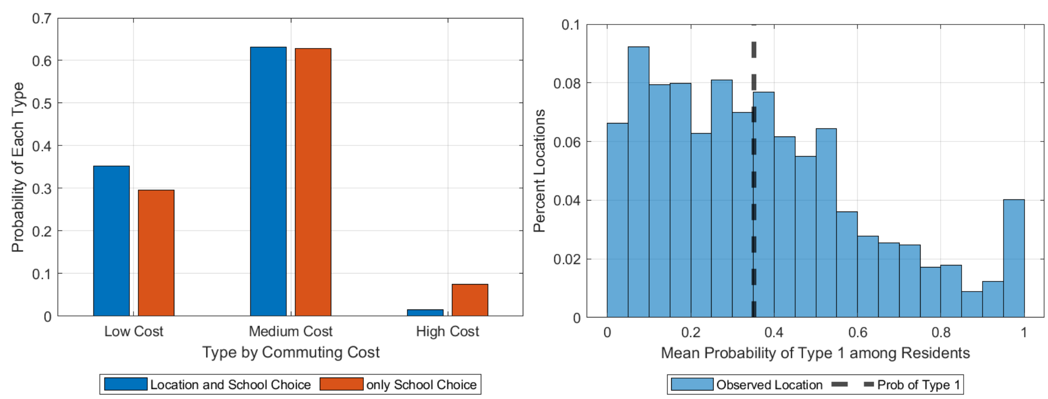
Table 1.3: Willingness to Pay

Next, we present households' WTP for a reduction in commuting time to school. Commuting time has been used as the numéraire in previous studies on public-school choice (Agarwal and Somaini, 2018a), and we convert it into the monetary term using housing cost. A median household is willing to pay $19 (=384/20 days) per school day to reduce commuting to school by 50 minutes a day.

We view our WTP estimate to capture various challenges that middle school students face during school commuting. For example, parents answer a survey by Sattin-Bajaj and Jennings (2022) that safety on the journey to a school is a main consideration factor for school application. Such concern of parents arises because many students commute to schools by themselves by public transportation or on foot. Middle school students are eligible for school bus service only in their first

estimate for a similar scenario is $1,740. ($= 0.7 \times 704 \times 12 \times (0.1/0.34)$), with a 0.7 adjustment to 2000 dollars using CPI (source: BLS CPI New York-Newark-Jersey City area) year if their schools offer any. Moreover, we calculate from the 2017 National Household Travel Survey that at least 70% of students in our sample commute to schools without any adult accompanied.[30] Finally, our WTP estimate is high to be interpreted as the forgone earning of middle school students. Adult commuters' value of commuting time is known to be 50%-70% of their hourly wage (Parry and Small, 2009; Purevjav, 2022), and the minimum wage in New York in 2015 was $9.

Overestimation of Commuting Cost Commuting cost is overestimated when we ignore households' residential sorting. We estimate a different version of the model without location choice (estimates are presented in subsection 1.9.4), and find that commuting cost is overestimated by 15% on average (mean β_i^d is -0.97 in a model with both location and school choice, and -1.11 with only school choice). Figure 1.6 describes what leads the model without endogenous location choice to overestimate commuting cost.



(a) Mean Probability of Unobserved Type (b) Location Sorting on Unobserved Type

Figure 1.6: Residential Sorting on γ and the Overestimation of Commuting Cost

Note: Panel (a) presents the probability of each unobserved type under each model. Panel (b) shows the mean probability of residents' being type 1 across residential locations.

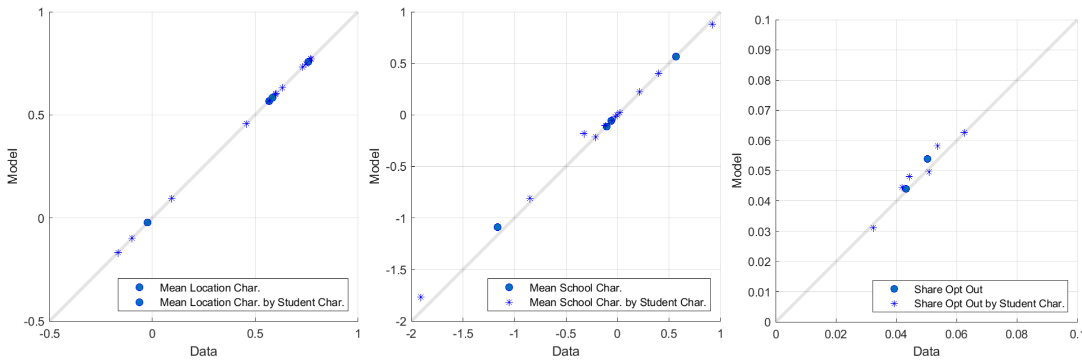
[30] To be accurate, 70% of middle school students residing in the NY-NJ-PA area, which is the finest geography available, commute to schools without any adult accompanied.

Panel (a) shows probabilities of types (q_k) from a model with both location and school choice compared with those from a model with only school choice. In the latter, we over-classify students into the high commuting-cost type, since we rationalize households' applying to schools nearby as due only to high commuting costs, as opposed to households' residential sorting based on unobserved school taste γ_i. Panel (b) plots the mean probability of being type 1 among residents across locations. For each student i, we calculate the probability of her being type 1 based on how well her location and school choices can be justified by being type 1 relative to other types (Bayes' rule). In the absence of residential sorting based on unobserved type, the mean probability of being type 1 among residents of a location should be similar across all locations. In contrast to this, some locations have zero type-1 students while others have many type-1 students, which implies sorting based on unobserved type.[31]

Model Fit We simulate choices using our estimates to validate whether our model can replicate the data patterns. To minimize the idiosyncrasies coming from preference shocks and the lottery number, we present the average over 100 simulations.

Figure 1.7 plots the simulated and observed moments from school and location choice. Moments include the mean observable characteristics of chosen options and the correlation between students' characteristics and those of their chosen options. Unsurprisingly, our targeted moments from location choice are well aligned with the 45-degree line. Meanwhile, even though we do not target school choice moments, and rather estimate school preference parameters via MLE, our simulated moments of school choices are close to data moments. Table 1.20 presents the numbers plotted in Figure 1.7.

[31]There is also an idiosyncrasy coming from a finite sample. Figure 1.15 compares the distribution from the data and that from a simulation in which we allocate households randomly across locations.



(a) Location Choice (b) School Choice (c) Enrollment

Figure 1.7: Model Fit

Note: We take the average over 100 simulations with draws of η, ϵ, and the lottery number. Moments include the mean observable characteristics of chosen options and the correlation between students' characteristics and those of their chosen options. In panel (b), we focus on students' first choice. In panel (c), we present the fraction of students who choose each outside option.

1.6 Source of School Segregation

In this section, we use model estimates to identify the sources of school segregation. Even with an extensive school choice system in place, NYC middle schools are highly segregated.[32] There are also large differences in academic achievement across these segregated schools. In the 2014-15 academic year, classmates of minority students (in their assigned schools) had standardized test scores than were one standard-deviation lower than the classmates of non-minority students. In this section, we explore which components of the model explain the cross-racial gap in the test scores of students' peers in their assigned schools

In Table 1.4, we investigate to what extent the cross-racial gap in test scores is explained by the following components of the model: access-to-school preference, heterogeneity in preference over location characteristics, and that over school characteristics. Column (1) in Table 1.4 presents the cross-racial gap under the status

[32]In terms of racial composition, 77% of Black and Hispanic students attend schools that enroll less than 10% of White students, while only 11% of White students and 43% of Asian students attend schools that enroll less than 10% of White students (Cohen, 2021).

quo; minority students attend schools with lower test scores by one student-level standard deviation than their non-minority peers.

Status Quo Racial Gap	(1)	(2)	(3)
	Racial Gap Explained by:		
	Access-to-school Preference	Heterogeneous Preference over: Other Location Characteristics	School Characteristics
-1.048	-0.312	-0.466	-0.182

Note: The cross-racial gap is the difference in test scores of the schools students attend for minority and non-minority students. We shut down each channel for one household one at a time. In column (2), we impose $\alpha^u = 0$. In column (3), we impose $\alpha^{WZ} = \alpha^{pZ} = 0$, In column (4), we impose $\beta^{XZ} = \beta^{dZ} = 0$.

Table 1.4: cross-racial Gap in Coassigned Peers' Test Score

Next, columns (2) and (3) demonstrate that residential sorting is the main driver of school segregation, which is in line with past studies (Laverde, 2020; Monarrez, 2020). We further break down what part of the gap is explained by residential sorting based on access to school (column (2)) versus sorting based on other location amenities (column (3)). Column (2) demonstrates that 31% of the gap observed in the data is explained by residential sorting based on access-to-school utility. In this scenario, we shut down residential sorting based on access to school ($\alpha^u = 0$). Thus, households choose locations as if they do not know that the locations chosen determine commuting costs and location-based admissions probabilities to schools. The cross-racial gap in this scenario comes from households' heterogeneous preferences over location characteristics other than access to school and those over school characteristics.

In columns (3) and (4), we investigate the role of preference heterogeneity. Column (3) shows that households' heterogeneous location preferences play a key role in generating school sorting. We shut down heterogeneous preferences over location characteristics and price by setting $\alpha^{WZ} = \alpha^{pZ} = 0$; thus residential sorting is only based on access to school. This scenario explains 46% of the gap.[33]

[33] This largely comes from heterogeneous preferences over location characteristics rather than price. Shutting down only the heterogeneous preference over housing costs reduces the gap by only 0.009.

In column (4), we shut down heterogeneous preferences over school characteristics by setting $\beta^{XZ} = \beta^{dZ} = 0$, so that households choose locations and schools as if they have perfect consensus over what makes a school desirable, even though they disagree on what makes a location desirable. This explains 18% of the cross-racial gap.

1.7 Citywide Access to Highest-achieving Schools

NYC middle schools are intensely segregated (Cohen, 2021; Idoux, 2022), and many believe that location-based priorities are the main cause. The city has long acknowledged this issue and proposed plans to relax location-based priorities, many of which have triggered heated debate among parents, students, and educators.[34]

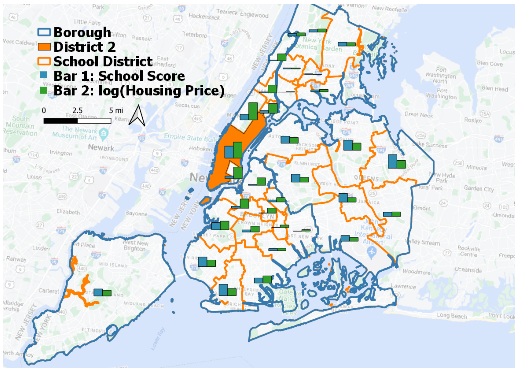
We evaluate a scenario in which we introduce purely lottery-based admissions to schools in School District 2. The district is located in lower Manhattan and has been at the center of ongoing policy debate regarding whether to retain location-based admissions rules.[35] In District 2, the average housing cost is about six times higher than other districts, and the mean test score for its schools is more than 1.2 standard deviations higher than those in other districts (Figure 1.8).

We compare the status quo, in which we simulate households' location and school choices under the current admissions rules, with the following scenarios in which we scrap all admissions criteria—both location-based priority rules and academic screening—for schools in District 2.

1. **OnlySC + No Opt-out**: Residential locations under status quo are fixed. We report the characteristics of the schools students are assigned to.

[34]For example, a plan to scrap all location-based priorities for high schools was canceled due to pushbacks from parents (Russo, Barbara , Zoned High School Options for NYC Students Will Remain in Place, NY Metro Parents, December 14 2021). Meanwhile, smaller plans have been implemented; For example, starting in the 2019-2020 academic year, Bronx middle schools have been open to all students in the Bronx (Zingmond, Laura, Bronx Middle School Best Tets, InsideSchools, October 20 2020).

[35]Shapiro, Eliza, N.Y.C. to Change Many Selective Schools to Address Segregation, the *New York Times*, December 18, 2020



(a) Location of District 2

	(1) District 2	(2) Others
Panel A: School Characteristics		
Mean z-score	1.124	-0.155
Safety	-0.087	-1.145
Share Minority	0.320	0.723
N of Schools	24	646
Panel B: Neighborhood Housing Price		
Unit Price (1K)	2,606	464
N of Neighborhoods	81	2,057

(b) District 2 Characteristics

Figure 1.8: District 2 Characteristics

Note: District 2 is the shaded area in the figure. School score is the mean of z-scores among enrolled students from the NYS standardized Math and Language test. Housing price is the mean price of residential units sold in 2013-14 located in each school district. Safety is a composite of crime incidences of different categories at the school building. Minorities include Black and Hispanic.

2. **LCSC + No Opt-out**: Households reoptimize residential locations. We report the characteristics of the schools students are assigned to.

3. **OnlySC + With Opt-out**: Residential locations under status quo are fixed. We report the characteristics of the schools students are enrolled in, excluding those who opt out.

4. **LCSC + With Opt-out**: Households reoptimize residential locations. We report the characteristics of the schools students are enrolled in, excluding those who opt out.

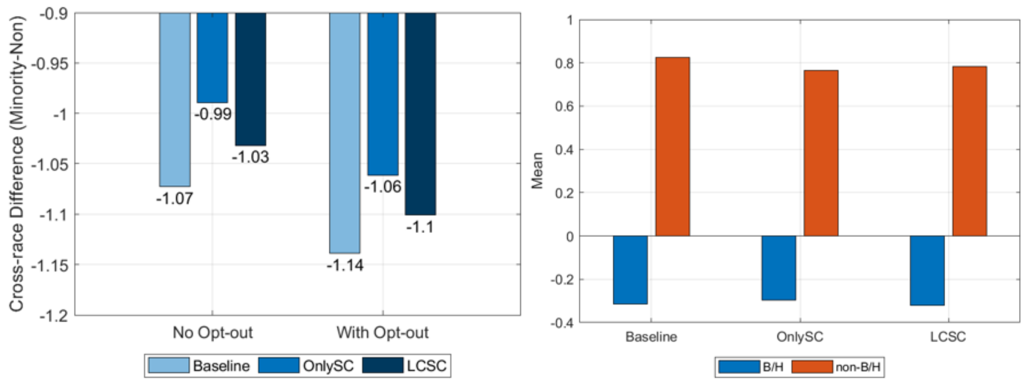
We solve the new equilibrium admissions cutoffs under the policy to address over-subscription to popular schools, especially District 2 schools. The main outcome of interest is the cross-racial gap in the characteristics of coassigned or co-enrolled school peers, which we interpret as the measure of inequity or school segregation. In **With Opt-out** cases, we calculate the mean characteristics of students who enroll in each school, excluding those who choose outside options.

We predict the effects of a policy that targets only one cohort of middle school applicants. Thus, we assume that the housing market can absorb changes in the demand of households with middle school applicants, who account for only 3% of the population. We also assume that school characteristics are invariant under a new policy. Furthermore, we compare the distribution of households across schools and residential locations in a steady state, since we do not model moving costs.

Cross-racial Gap in Peer Characteristics Figure 1.9 shows the gap in coassigned or coenrolled peers' test scores between minority and non-minority students in each scenario. While the reform narrows the cross-racial gap in peer test scores, households' location choices dampen such effect. The y-axis in panel (a) is the cross-racial difference school peers' standardized test scores. The policy closes the cross-racial gap in coassigned peers' test scores from 1.07 to 0.99, thus approximately 7%, if households' residential locations were fixed (**No Opt-out, OnlySC**).[36] However, when households reshuffle across locations, the effect reduces to 3.3% (**No Opt-out, LCSC**). The cross-racial gap in assigned schools (**No Opt-out**) is always smaller than that in enrolled schools (**With Opt-out**). But the effects of the policy and households' endogenous location choices on the cross-racial gap in enrolled schools are similar to those on the cross-racial gap in assigned schools. In panel (b), we present the mean of coenrolled peers' test scores by minority and non-minority students. It shows that the policy closes the cross-racial gap both because non-minority students enroll with lower-achieving peers and minority students enroll with higher-achieving peers.

Location Choice Patterns Next, we delve into households' location choices to understand how those dampen the equity impact of the policy. The key is that locations decide on commuting costs as well as location-based priorities, which together determine access-to-school utility of locations. Under the status quo,

[36]Zooming in District 2 schools, the cross-racial gap reduces from 0.35 standard deviation to 0.15 standard deviation, thus the gap reduces by 57% among students who are assigned to District 2 schools.



(a) Gap in Peers' Test Score (1 std) (b) Mean of Peers' Test Score (1 std)

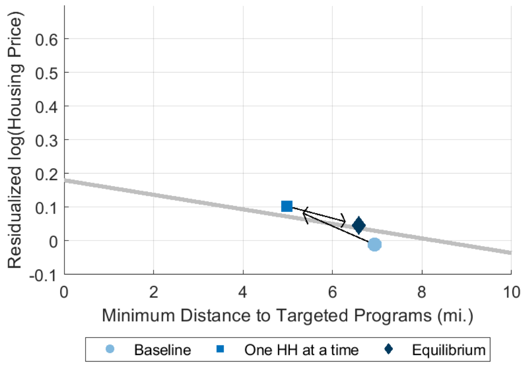
Figure 1.9: Cross-racial Gap in School Characteristics

Note: Panel (a) shows the difference in mean test scores of coassigned/coenrolled peers between Black/Hispanic and other students. Panel (b) shows the mean test score of coenrolled peers for each group separately. We use z-scores from the NYS standardized Math and Language test.

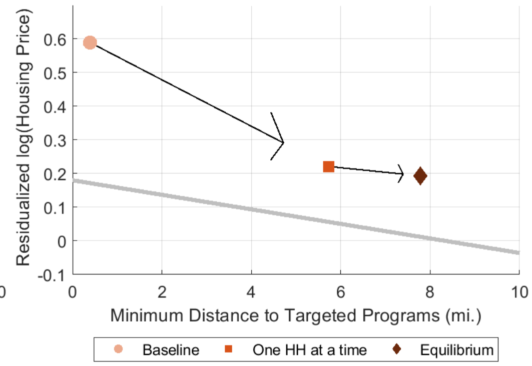
households have positive admission chances to District 2 schools only when they reside in District 2. Hence, utility from access to District 2 schools differs only by whether a location is either within or outside the district. On the other hand, the policy equalizes admissions probability to District 2 schools across locations. Hence, locations differ in utility from access to District 2 schools by their proximity to District 2. Standing in contrast to the status quo, locations outside District 2 have different levels of utility from access to District 2 schools from one another.

These changes in access-to-school utility result in a different reoptimization in location choice patterns among households who live in District 2 under the status quo (=*D2 residents*) and others (=*Non-D2 residents*). We first present the location choice patterns of these two groups in a scenario in which we introduce purely lottery-based admissions to District 2 schools to *one household at a time*. In this scenario of *one household at a time*, a given household does not expect other households to modify their behavior in response to the policy change.

In Figure 1.10, the y-axis is the demeaned log of housing price and the x-axis is the average distance to schools in District 2 from each location. Each dot describes the mean characteristics of locations chosen by *Non-D2 residents* (panel (a)) and



(a) Non-D2 Residents



(b) D2 Residents

Figure 1.10: Location Choice Patterns

Note: Panel (a) illustrates the location choices of *Non-D2 residents*, and (b) of *D2 residents*. The x-axis is the average distance to schools in District 2 from locations. The y-axis is the residualized log sales price. Each dot shows the median characteristics of locations chosen by *Non-D2 residents* and *D2 residents* who change locations under the policy, respectively. In each panel, we plot location choice patterns when we grant citywide access to District 2 schools to one household at a time and when we grant citywide access to all households. For the latter, we solve equilibrium admissions cutoffs.

D2 residents (panel (b)).

In the *one household at a time* scenario, *Non-D2 residents* relocate closer to District 2 at the expense of higher housing costs (panel (a)). While the policy makes them eligible to apply to and enroll in District 2 schools, such an option is not attractive when they stay in their baseline locations due to the high commuting cost to District 2 schools. Meanwhile, *D2 residents* choose locations with lower housing costs, but farther from District 2 schools (panel (b)). Purely lottery-based admissions make it no longer necessary to live in District 2 to ensure positive admissions probabilities to District 2 schools.

In equilibrium, the location choice behaviors of *Non-D2 residents* are largely muted, while those of *D2 residents* are reinforced. This is because citywide access to District 2 schools induces applicants from a broader area, and thus the admissions chances to District 2 schools are lower from each household's point of view. This makes choosing locations nearer to District 2 by *Non-D2 residents* less attractive and choosing locations farther from District 2 by *D2 residents* more attractive.

Connection between Location Choice and Peer Characteristics Table 1.5 reveals the link between households' location choice and the school desegregation effect of the policy. *Non-D2 residents*' spatial reshuffling narrows the cross-racial gap in school characteristics. For example, by relocating, minority *Non-D2 residents* are assigned to schools with a 13.7-percentage-point lower minority share. This largely comes from their choosing locations nearer to District 2 schools and more actively applying to and enrolling in those schools.

	(1)	(2)	(3)	(4)
	Non-D2 Resident		D2 Resident	
	Non-Minority	Minority	Non-Minority	Minority
Share	33.53%	62.81%	2.72%	0.94%
Panel A: Location Choice when Granting Citywide Access to One HH at a Time				
Change Location under New Policy?	0.220	0.176	0.592	0.511
Conditional on Changing Location:				
Δ Frac. Minority of Assigned School	0.107	-0.137	-0.150	-0.085
Δ Mean Score of Assigned School	-0.204	0.181	0.273	0.095
Δ Frac. Minority of Neighborhood	0.047	-0.216	0.147	0.344
Panel B: Location Choice in Equilibrium				
Change Location under New Policy?	0.136	0.092	0.974	0.936
Conditional on Changing Location:				
Δ Frac. Minority of Assigned School	0.022	-0.048	-0.186	-0.023
Δ Mean Score of Assigned School	-0.211	0.030	0.281	-0.099
Δ Frac. Minority of Neighborhood	0.004	-0.042	0.130	0.416

Note: Minority includes Black or Hispanic. *D2 residents* are those who reside in one of the locations in District 2 under the status quo. Each column shows the mean of variables for each group.

Table 1.5: From Location Choice To School Assignment

The location choice patterns of *D2 residents* stand in contrast to those of *Non-D2 residents*; their spatial reshuffling dampens the equity impact of the policy. They seek locations that come with a secured seat in higher-achieving and lower-minority schools, and the purely lottery-based admissions to District 2 schools make locations within District 2 less attractive. Instead, they choose locations where location-based admissions are kept in place. By doing so, they are assigned to schools with a 15 percentage point lower minority share.

Previously, we have shown while relocation motives of *Non-D2 residents* are muted in equilibrium those of *D2 residents* are reinforced (Figure 1.10). Then, Table 1.5 shows relocation of *Non-D2 residents* amplifies the equity impact of the policy, while that of *D2 residents* dampens the impact. Combined together, endogenous location choices dampen the effect of the policy on the cross-racial test score gap, as depicted in Figure 1.9.

Commuting Distance to School and Welfare Another widespread concern about relaxing the importance of location-based admissions priority is that students would have to commute longer distances. In Table 1.6, we present average commuting distances to schools under each scenario by minority and non-minority students. We also present the change in welfare, which is a number that summarizes various changes in outcome induced by the counterfactual policy.[37]

	(1)	(2)	(4)	(5)	(6)	(7)	(8)	(9)
	Distance to school		Δ Welfare (% Housing Cost)					
	Non-Minority	Minority	Non-Minority	Minority	Overall			
	Mean	Mean	Mean	Mean	Mean	p50	Sum Loss	Sum Gain
Baseline	1.450	1.595	-	-	-	-	-	-
OnlySC	1.728	1.645	-0.074	-0.006	-0.030	0.012	894.620	591.450
LCSC	1.738	1.769	0.005	0.045	0.031	0.029	264.270	572.680

Note: Minority includes Black or Hispanic. *D2 residents* are those who reside in one of the locations in District 2 under the status quo. The fastest driving distance between a school and a Census block is calculated using Open Route Services. Welfare is measured by exante utility (Equation 1.10 at the chosen location), which we we convert into log housing cost. We present the difference in welfare under the policy relative to the baseline scenario.

Table 1.6: Effect on Commuting Distance and Welfare

Commuting distances increase for both minority and non-minority students, so the gap is decided by which group experiences a larger increase. While the policy narrows the cross-racial gap in commuting distance under **OnlySC**, households' endogenous location choice partially undoes this effect (**LCSC**).

[37]While the welfare measure is a good summary of various changes, we might want to use caution in interpreting this. This is because our demand estimates might not represent households' true preferences, even though they capture how households make location, school, and enrollment choices. Former studies have documented various types of friction in location and school choices such as limited information, limited attention, and even discrimination by landlords or schools (Christensen and Timmins, 2018; Luflade, 2018; Allende et al., 2019; Son, 2020; Christensen et al., 2020; Ferreira and Wong, 2020).

Next, we calculate the change in welfare, which we measure with the ex-ante utility (Equation 1.10) at each household's optimal location. We convert the utility into percentage housing cost for ease of interpretation. Since we restrict households from reoptimizing locations, **OnlySC** mechanically gives lower welfare in comparison to the other two scenarios.

Under **OnlySC**, both minority and non-minority students experience a decrease in welfare on average. This comes from a large decrease in welfare among *D2 residents*, who cause the distribution of welfare change to skewed to the left. Indeed, a median household experiences an increase in welfare by 1.2% of housing cost. The sum of welfare losses is greater than the sum of gains by 300% of housing cost. This suggests while the policy might be approved by the voting among these households, it might face harder pushback from households who lose.

In the long run, where households adjust their locations (**LCSC**), both the average and the median household experience welfare gain, by 3.1% and 2.9%, respectively. The benefit is largely concentrated among minority students (a 4.5% increase), who obtain eligibility to District 2 schools while living in affordable locations. We consider this welfare gain an upper bound given that we assume away from moving cost.

Other Margins The counterfactual policy has effects that go beyond changing the cross-racial test score gap, which we briefly discuss here. First, residential segregation, which we measure with an entropy-based segregation measure (subsection 1.9.5) decreases by 1.5% by race and 10% by income. Second, the policy increases the mean score of peers even among the lowest-performing minority students, an effect that is also dampened by households' location choices. There is a 6.5% increase in coassigned peers' test scores in **OnlySC** and a smaller effect (4% increase) in **LCSC**.

Other Policy Plans Next, we discuss the impact of another plan to introduce purely-lottery based admissions to District 26 schools. District 26 is located in upper Queens and features the highest mean test score of schools. The average housing unit price was $614,000, which is about a quarter of the housing price in District 2.

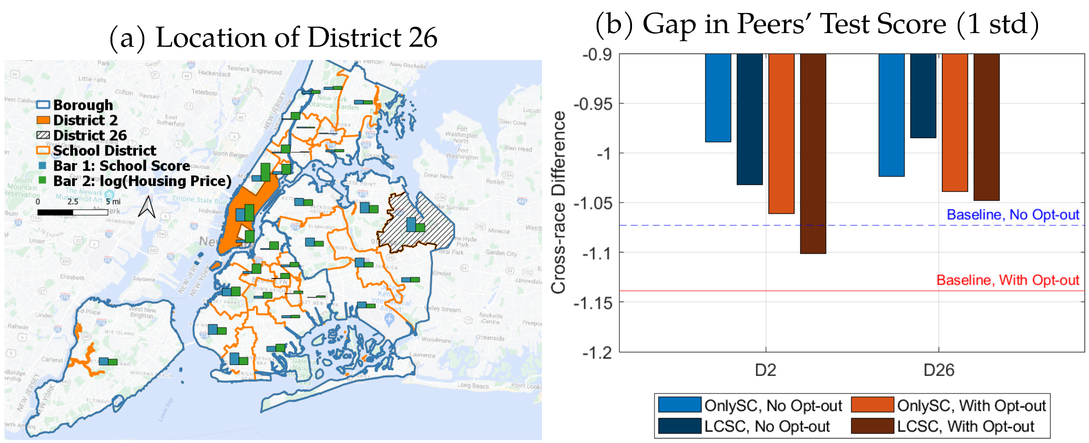


Figure 1.11: Lottery-based Admissions to District 26 Schools

Note: Panel (a) shows the location of District 26. Panel (b) shows the difference in mean test scores of coassigned/coenrolled peers between Black/Hispanic and other students. We use z-scores from the NYS standardized Math and Language test. Under **D2**, we introduce purely-lottery based admissions to schools in school District 2. Under **D26**, we target District 26.

Figure 1.11 presents the characteristics of District 26, and how the cross-racial gap changes across scenarios. First, whether and how households' endogenous location choices change the equity impact of the policy varies across policies. Focusing on **No Opt-out** scenarios, while endogenous location choices dampen the effect of the lottery-based admissions to District 2 schools by half, it amplifies the effect of the policy targeting District 26. This is because, in the latter scenario, minority households' endogenous location choice responses to shorten commuting distances to District 26 schools dominate non-minority households' location choices to get away from the policy. The lower housing cost of District 26 relative to District 2 is the key reason.

Second, the comparison across policies changes depending on if we consider households' endogenous location choice or not. Focusing on **No Opt-out** scenarios, while targeting District 2 schools seems more effective in reducing the cross-racial gap when we take residential locations as given, households' endogenous location choices in response make targeting District 26 more effective.

Lastly, households substitute between opting out to outside schooling options and reoptimizing their residential locations. While opt-out plays a minor role when we target District 2 schools, its role is more pronounced when we target District 26 schools. This is because households who live in District 26 cannot afford all other school districts as District 2 do. Thus, they take advantage of outside options to enroll their kids in a more preferable school when they lose the advantage in admissions changes to District 26 schools.

We also present results when we scrap location-based admissions to District 2 schools but with academic screening in place in Figure 1.16. The policy does not close the cross-racial gap in school peers' test scores when households' residential locations are fixed. Instead, households' spatial reshuffling rather widens the gap by 2.6%. This is because non-minority, higher-achieving students who reside outside of District 2 are largely motivated to move locations nearby District 2, which pushes minority students in District 2 out to schools with a higher proportion of minority peers.

1.8 Conclusion

Increasingly more school districts have adopted centralized school choice systems in the hope that they can break the tie between spatial disparities and educational disparities. Whether they can achieve these goals, however, crucially depends on the extent to which students are willing to take advantage of school choice options as well as how households respond to the policy by (1) reshuffling across locations and (2) opt-out to other schooling options.

This paper develops a unified framework of households' residential location choice and school choice under a centralized school choice system. By doing so, we extend empirical school choice literature that has studied many factors for students' school applications and assignments but has given little attention to endogenous residential location choices. Residential locations determine location-based admissions probabilities and commuting distances to schools, which motivates households to choose locations by considering such ties. Our framework captures this as well as the possibility of opting out to outside schooling options. Rich heterogeneity in households' observed and unobserved preferences over various school and location characteristics generates sorting into locations and schools. We map the framework to New York City's middle school choice context, which is the largest unified district with a centralized school choice system.

Our policy analysis shows how a radical school desegregation effort might have a minimal effect, largely because of households' choosing locations that can undo the policy. The policy grants citywide access to the school district that covers lower- and mid-Manhattan. We find that households' spatial reshuffling dampens the policy effect by half. Some minority households choose locations from which the commute to top district schools is easier, which amplifies the desegregation effect. However, other non-minority households choose locations that come with secured seats in higher-achieving schools outside of the affected district, which undoes the effect of the policy.

Several lines of inquiry are left for future work. First, such work might quantify the complementary effect between school desegregation policies and housing market policies. We find that 45% of school segregation is explained by households' heterogeneous preference over location characteristics other than price and access to school. Recent evidence shows that such heterogeneity stems from information frictions (Ellen et al., 2016; Ferreira and Wong, 2020) or housing market discrimination (Christensen and Timmins, 2018), which suggests that policy interventions can change how households choose locations. Second, although we take the location of schools as given in the paper, future work can consider where to open a new school or how to allocate resources to schools in different locations. Such work informs policymakers' ongoing efforts to design school choice systems that could benefit a larger number of students.

1.9 Appendix

1.9.1 Details of NYC School Choice Process

Student-Proposing Deferred Acceptance Algorithm In detail, DA works as follows (Gale and Shapley, 1962; Abdulkadiroğlu et al., 2005):

- **Step 1**
  Each student proposes to her first choice. Each program tentatively assigns seats to its proposers one at a time, following their priority order. The student is rejected if no seats are available at the time of consideration.

- **Step $k \geq 2$**
  Each student who was rejected in the previous step proposes to her next best choice. Each program considers the students it has tentatively assigned together with its new proposers and tentatively assigns its seats to these students one at a time following the program's priority order. The student is rejected if no seats are available when she is considered.

- The algorithm terminates either when there are no new proposals or equally when all rejected students have exhausted their preference lists.

DA produces the student-optimal stable matching and is strategy-proof i.e., truth-telling is a weakly dominant strategy for students.

NYC School Admission Methods Middle school programs use a variety of admission methods—Unscreened, Limited Unscreened, Screened, Screened: Language, Zoned and Talent Test. Unscreened programs admit students by a random lottery number, and Limited Unscreened programs use rules that give priority to those who attend information sessions or open houses. Screened programs as well as Screened: Language programs select students by program-specific measures such as elementary school GPA, statewide test scores, punctuality and interviews. Zoned programs guarantee admissions or give priority to students who reside in the school's zone, and Talent Test programs use auditions.

The Timeline of Admission Process The timeline of the admission process is as follows (Corcoran and Levin (2011a), Directory of NYC Public High Schools). By December, students are required to submit their ROLs. By March, DA algorithms are run and determine students' assignments. Students who accept their offer finalize, and if a student rejects an offer, then she goes to the next round. This describes the main round of the entire system. A majority of students finalize in the main round (about 90% each year). Students who are not assigned in the main round or rejected the assignment go to the Supplementary round which is similarly organized as the main round and includes programs that did not fill up their capacities in the main round, or programs that are newly opened. Finally, there is an administrative round in which students who are not assigned a school even after the second round are administratively assigned to a school.

Example of ROL Figure 1.12 shows an example of rank ordered list.

Rank your choices in order of preference here.

Choice Number	District	Program Code	School Name/Program Name
2	1	M378L	School for Global Leaders
3	21	K239CM	Mark Twain (I.S. 239) Magnet Program - Computer/Math
1	21	K239VO	Mark Twain (I.S. 239) Magnet Program - Vocal
	1	M292S	Henry Street School for International Studies
4	1	M140S	The Nathan Straus Preparatory School of Humanities (P.S. 140)
	1	M301S	Technology, Arts, and Sciences Studio

SAMPLE

Source: NYC DOE Middle School Directory 2014-15

Figure 1.12: Example of Customized List and Rank-Ordered List

1.9.2 Supplementary Materials for section 1.3

Cleaning Procedure of DOF Annualized Selling Record First, we drop non-residential properties such as industrial buildings, commercial buildings, and vacant land, based on both the tax class and building class. Then we merge the selling record with the Primary Land Use Tax Lot Output (PLUTO) to recover the exact location of each sold property.[38] Lastly, we exclude transactions that are unlikely at arm's-length. We drop transactions of zero price that include transfers within a family. Also, we drop records of significantly low prices relative to other properties of similar characteristics. Specifically, we run a hedonic pricing model that includes tax class, assessment value, the interaction of the two, calendar time FE at each borough, month FE by borough, building type, land area, building area, total unit, odd shape, age, age square, garage area, the year of alteration, and commercial area (R square = 0.67). Then we drop observations of the predicted residual is less than 1 percentile. We run the regression separately for coops. Following Schwartz et al. (2014), we lag housing cost by one year to take into account that there could be some time lag for school quality to be capitalized in the housing cost.

Housing Cost and Structure in ACS 5-year Estimates While ACS 5-year estimates capture the price and characteristics of representative housings, the biggest limitation is that each observation is at the Census block group level, which could be too coarse to capture the change within a narrow bandwidth around the boundary. 1,944 of 7,506 Census block groups whose centroid is within 0.2 miles from a school district boundary overlay across a school district boundary.[39] Thus, the distance from the centroid of a Census block group to the closest boundary is a crude measure of proximity to the boundary.

Therefore, we further exploit the variation of population density across Census blocks within a block group. Table 1.7 illustrates two exemplary cases. Consider two census block groups A and B whose distances from their centroid to the closest boundaries are the same. Census block group A consists of two Census blocks, one of which is 0.15 miles away from the boundary and the other 0.28 miles away. Note that the two Census blocks differ in population density. Out of 90 occupied

[38]Two data sets are merged based on the identifiable tax lot number. One complication is in merging condos. The selling record has a unique id for each unit, while PLUTO for each condo. We use the Department of City Planning Property Address Directory that lists unit ids to a matching condo id.

[39]On the contrary, only 489 out of 38,498 Census blocks overlay across a boundary.

Census Block Group A:	Block A-1	Block A-2
Distance to Boundary (mi.)	0.15	0.28
The Number of Occupied Units	30	60
Census Block Group B:	Block B-1	Block B-2
Distance to Boundary (mi.)	0.15	0.28
The Number of Occupied Units	60	30

Note: Consider two Census Block Groups A and B with the same distance from their centroids to the nearest boundaries.

Table 1.7: Example of Census Block Groups

units in the block group, two-thirds are living in A-2. Census block group B has the opposite pattern.

To consider such differences in density, we weigh Census block groups with the percent of occupied units in Census blocks within 0.25 miles from the closest boundary when running Equation 1.1. In the example, Census block group A is given a weight of 0.33, while B has a weight of 0.66. We present the estimated effects of district school quality in Table 1.14.

Evidence Supporting the Identification Assumption of BDD The identification assumption of a boundary discontinuity design is that unobserved location amenities are as good as random within a narrow buffer around a boundary. While we cannot check this assumption directly, we present that other observed location characteristics are continuous in geography, which suggests that the assumption is plausible in this context.

Table 1.15 reports estimates $\hat{\beta}$ (Equation 1.1) for various housing characteristics and urban amenities. $\hat{\beta}$s for most of the variables are not statistically significant. One exception is that sold properties located within a school district with higher school quality are more likely to have been renovated (p-value < 0.05). Thus, we use residualized prices when estimating the model to absorb variations coming from housing characteristics and urban amenities, including the renovation status.

While each variable is not the main driver of sharp change in sales price at boundaries (Figure 1.3 and Table 1.10), a *set* of variables might. Table 1.16 checks this further. First, we run a hedonic regression of log sales prices on various housing characteristics and urban amenities using transaction records within a 0.25-mile buffer around boundaries. Then we sum variables using coefficients from the hedonic regression, which capture the extent to which each variable explains the variation in sales prices. Finally, we run the BDD regression (Equation 1.1) using the predicted prices as dependent variables and check if estimates $\hat{\beta}$ are significant.

Columns (1)-(2) of Table 1.16 describe that even a very extensive set of housing characteristics and urban amenities does not explain the sharp change in sales prices at boundaries—i.e., change in school quality is the main driver. Meanwhile, the estimate $\hat{\beta}$ is marginally significant (p-value < 0.1) when we include neighbors' composition (column (3)), which captures households' residential sorting at boundaries as well as their preference over neighbors' composition.

Supplementary Tables

Variables	(1)	(2)	(3)	(4)
	All		< 0.25-mile Buffer	
	mean	std	mean	std
Panel A: Student Characteristics				
Asian	0.185	0.388	0.162	0.368
Black	0.295	0.456	0.333	0.471
Hispanic	0.380	0.485	0.407	0.491
White	0.125	0.331	0.085	0.279
FRL	0.736	0.441	0.783	0.413
Standardized English Score	0.028	1.009	-0.077	0.997
Standardized Math Score	0.172	0.979	0.074	0.963
N	57,593		18,761	
Panel B: Sold Properties' Characteristics				
Unit Price ($1000)	802.8	2,063	707.9	1,496
Age	70.88	32.68	80.71	97.16
Number of Floors	7.150	9.083	6.493	8.105
Coop[a]	0.299	0.458	0.287	0.452
Manhattan	0.245	0.430	0.170	0.376
Bronx	0.079	0.270	0.088	0.283
Brooklyn	0.263	0.440	0.441	0.497
Queens	0.319	0.466	0.300	0.458
Staten Island	0.0934	0.291	0	0
N	106,040		23,836	
Panel C: Census Block Group Characteristics				
Median Rent	1,404	503.4	1,255	501.5
Median Value ($1000)	636.7	355.9	612.9	331.1
Median Age	70.83	13.45	72.67	13.08
% College and Higher Degree	0.352	0.237	0.298	0.229
% Minority	0.272	0.316	0.334	0.309
N	4,828		609	

Note: Source of each data set is NYC Department of Education, NYC Department of Finance Selling Record, and ACS 5-year estimates. All from 2013 to 2017.

Table 1.8: Summary Statistics: All vs. Sample for BDD

[a] We control for properties' co-op status in our analyses. Coops take up a large proportion of the NYC housing market (35% of sold properties, 50% of the housing stock) with two unique features. First, they are more common in Manhattan compared to other boroughs, and second, are cheaper to buy but come with high monthly maintenance fees. (Susan Stellin, Co-op vs. Condo: The Differences Are Narrowing, The New York Times, Oct. 5, 2012) Ignoring the composition of co-op and other housing types understate housing cost in Manhattan because the Sales files cover only sold price.

	(1)	(2)	(3)	(4)
Bandwidth	< 0.25 mile		< 0.2 mile	
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Student Characteristics	No	Yes	No	Yes
District School Quality	0.360	0.270	0.324	0.256
	(0.074)	(0.053)	(0.072)	(0.054)
N	16576	15809	13261	12657
R2	0.249	0.394	0.245	0.389
$\bar{y}$	-0.078	-0.055	-0.092	-0.068
$\mathtt{std}(y)$	0.800	0.779	0.796	0.773

Note: The dependent variable is the mean score of the schools middle-school-applying residents in a Census block are assigned to. Sample of 5th-grade students in academic year 2014-15 living in Census blocks within a buffer from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries.

Table 1.9: Effects of District School Quality on Assigned Schools' Quality

	(1)	(2)	(3)	(4)
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes
Urban Amenities	No	No	No	Yes
Panel A: 0.25-mile Buffer				
District School Quality	0.181	0.204	0.101	0.102
	(0.061)	(0.044)	(0.040)	(0.040)
N	23786	23786	23786	23786
R2	0.409	0.489	0.505	0.505
$\bar{y}$		12.88		
$\mathtt{std}(y)$		1.112		
Panel B: 0.2-mile Buffer				
District School Quality	0.108	0.150	0.073	0.073
	(0.068)	(0.049)	(0.045)	(0.045)
N	19057	19057	19057	19057
R2	0.401	0.480	0.493	0.494
$\bar{y}$		12.84		
$\mathtt{std}(y)$		1.086		

Note: The dependent variable is the log sales price of a residential unit. Sample of residential units sold within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table 1.10: Effects of District School Quality on Housing Sales Price

	(1)	(2)	(3)	(4)	(5)
Boundary FEs	Yes	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes	Yes
Urban Amenities	No	No	No	Yes	Yes
N of Population	No	No	No	No	Yes
Panel A: 0.25-mile Buffer					
District School Quality	0.183	0.266	0.805	0.789	0.808
	(0.176)	(0.175)	(0.222)	(0.225)	(0.219)
N	3755	3755	3755	3755	3755
R2	0.227	0.251	0.308	0.318	0.331
$\bar{y}$			3.515		
$\mathtt{std}(y)$			4.122		
Panel B: 0.2-mile Buffer					
District School Quality	0.216	0.280	0.766	0.750	0.830
	(0.180)	(0.183)	(0.240)	(0.244)	(0.242)
N	2970	2970	2970	2970	2970
R2	0.223	0.246	0.295	0.303	0.313
$\bar{y}$			3.490		
$\mathtt{std}(y)$			4.033		

Note: The dependent variable is the number of middle-school-applying residents in a Census block. Sample of Census blocks within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in panel A and the 0.2-mile buffer in panel B. Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block. The number of population is at Census block group level, which we obtain from the ACS 5-year estimate.

Table 1.11: Effects of District School Quality on the Number of Middle-school-applying Residents

	(1)	(2)	(3)	(4)
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes
Urban Amenities	No	No	No	Yes
Panel A: 0.25-mile Buffer				
District School Quality	-0.139	-0.143	-0.065	-0.065
	(0.029)	(0.028)	(0.016)	(0.016)
N	2970	2970	2970	2970
R2	0.480	0.496	0.595	0.596
$\bar{y}$		0.620		
$\mathtt{std}(y)$		0.417		
Panel B: 0.2-mile Buffer				
District School Quality	-0.132	-0.132	-0.066	-0.067
	(0.034)	(0.033)	(0.024)	(0.024)
N	2353	2353	2353	2353
R2	0.484	0.501	0.603	0.603
$\bar{y}$		0.633		
$\mathtt{std}(y)$		0.415		

Note: The dependent variable is the share of Black and Hispanic applicants among middle-school-applying residents in a Census block. Sample of Census blocks within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table 1.12: Effects of District School Quality on Minority Share of Middle-school-applying Residents

	(1)	(2)	(3)	(4)
Boundary FEs	Yes	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes	Yes
Urban Amenities	No	No	No	Yes
Panel A: 0.25-mile Buffer				
District School Quality	-0.139	-0.142	-0.091	-0.094
	(0.038)	(0.034)	(0.033)	(0.033)
N	2970	2970	2970	2970
R2	0.203	0.218	0.261	0.263
$\bar{y}$		0.683		
$\mathtt{std}(y)$		0.357		
Panel B: 0.2-mile Buffer				
District School Quality	-0.139	-0.141	-0.099	-0.102
	(0.042)	(0.038)	(0.037)	(0.037)
N	2353	2353	2353	2353
R2	0.173	0.192	0.238	0.240
$\bar{y}$		0.698		
$\mathtt{std}(y)$		0.350		

Note: The dependent variable is the share of free or reduced lunch eligible applicants among middle-school-applying residents in a Census block. Sample of Census blocks within a bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. We use the 0.25-mile buffer in columns (1)-(2) and the 0.2-mile buffer in columns (3)-(4). Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table 1.13: Effects of District School Quality on Poverty Share of Middle-school-applying Residents

	(1)	(2)	(3)
Boundary FEs	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes
Housing Characteristics	No	Yes	Yes
Neighborhood Characteristics	No	No	Yes
Urban Amenities	No	No	No
Panel A: Log(Median Gross Rent)			
District School Quality	0.163	0.159	0.058
	(0.047)	(0.047)	(0.035)
N	1875	1873	1873
R2	0.352	0.374	0.535
$\bar{y}$		7.119	
$\mathtt{std}(y)$		0.377	
Panel B: Log(House Value)			
District School Quality	0.057	0.067	0.058
	(0.066)	(0.061)	(0.062)
N	1332	1331	1331
R2	0.478	0.540	0.558
$\bar{y}$		13.20	
$\mathtt{std}(y)$		0.540	

Note: The dependent variable is the log median gross rent in panel A, and the log house value reported by homw owners in panel B. Unit of analysis is Census block group, where we weigh block groups by the share of occupied units in Census blocks within a 0.25-mile buffer from the closest school district boundary (See Section 1.9.2). District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table 1.14: Effects of District School Quality on Rent and House Value

	(1)	(2)	(3)	(4)
Panel A: Housing Characteristics of Sold Properties				
Dependent Variable:	N of Floors	Coop	Commercial Area (1K sqft)	Renovation
District School Quality	-1.55	0.076	-5.84	0.100
	(1.25)	(0.047)	(5.53)	(0.038)
N	23786	23786	23786	23786
R2	0.466	0.230	0.182	0.165
$\bar{y}$	6.498	0.288	11.19	0.809
$\mathtt{std}(y)$	8.112	0.453	56.27	0.393
Panel B: Urban Amenities				
Dependent Variable:	Bus Stop	Subway Station	Laundries	Café
District School Quality	-0.014	-0.004	0.003	-0.004
	(0.024)	(0.008)	(0.008)	(0.007)
N	8091	8091	32340	32340
R2	0.025	0.020	0.068	0.087
$\bar{y}$	0.127	0.019	0.052	0.014
$\mathtt{std}(y)$	0.413	0.140	0.252	0.157

Note: Sample of residential properties sold within a bandwidth from the closest school district boundary in Panel A. Sample of Census block groups whose centroids are within a bandwidth from the closest school district boundary in Panel B. Sample of Census blocks whose centroids are within a bandwidth from the closest school district boundary in Panel C. We use 0.25 mile buffer. In panel B, Each Census block groups is weighted accroding to the procedure explained in the appendix. Standard errors in parentheses are clustered at each tax lot, Census block group, and Census block, respectively.We allow the local cubic control of distance to differ at the opposite side of boundaries.

Table 1.15: Housing Characteristics and Urban Amenities at School District Boundaries

	(1)	(2)	(3)
Boundary FEs	Yes	Yes	Yes
Local Cubic Control for Distance	Yes	Yes	Yes
Housing Characteristics	Yes	Yes	Yes
Neighborhood Characteristics	No	No	Yes
Urban Amenities	No	Yes	Yes
Panel A: 0.25-mile Buffer			
District School Quality	-0.027	-0.018	0.094
	(0.069)	(0.065)	(0.052)
N	23786	23786	23786
R2	0.457	0.479	0.687
$\bar{y}$	12.88	12.88	12.88
$\mathtt{std}(y)$	0.546	0.560	0.728
Panel B: 0.2-mile Buffer			
District School Quality	-0.021	-0.002	0.069
	(0.077)	(0.074)	(0.057)
N	19057	19057	19057
R2	0.443	0.467	0.686
$\bar{y}$	12.84	12.84	12.84
$\mathtt{std}(y)$	0.509	0.526	0.696

Note: The dependent variable is the predicted log sales price of a residential unit, which we construct by running a hedonic regression of log sales price on covariates, as explained in Section 1.9.2. Sample of residential units sold within 0.25-mile bandwidth from the closest school district boundary. District school quality is measured by the mean NYS standardized test score of students enrolled in middle schools (previous cohorts) in the district. Standard errors in parentheses are clustered at school district level. The local cubic control of distance differs at the opposite side of boundaries. Housing characteristics include the space of the unit, land use of the tax lot, number of floors, age, renovation status, and storage area of the building, all of which we interact with a dummy if the property is coop. Neighbor characteristics include % minority, median household income, % college-or-more-educated, and median commuting time to work at Census block group. Urban amenities include the number of bus stops, subway stations, laundries, cafes, and crime incidents of different categories at Census block.

Table 1.16: Effects of District School Quality on Hedonic Sales Price

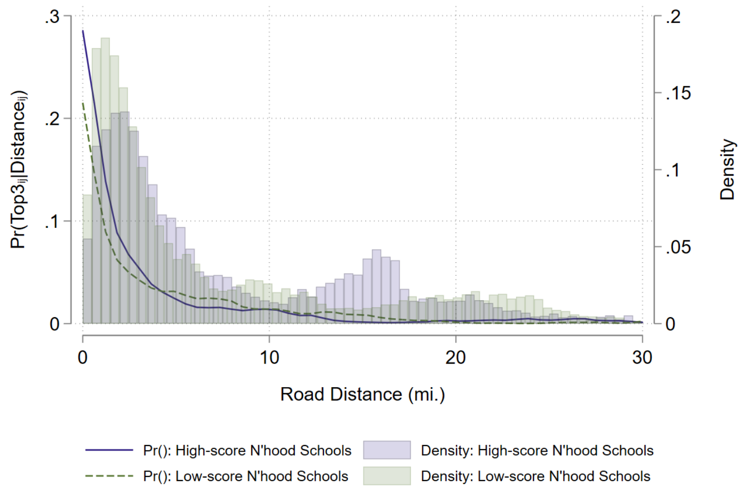


Figure 1.13: Probability of Listed as Top 3 Given Distance

Note: The fastest driving distance to school is calculated using the Open Route Service (ORS). the y-axis of the graph presents the probability of choosing school j as the top 3 choices among all schools a student is eligible for. Lines present the probability of student i's listing school j as one of her top 3 choices as a function of road distance between i and j, for all pairs of (i, j). We present the pattern separately for students whose neighborhood schools' mean test score is greater/smaller than the average, where we use the three closest schools as neighborhood schools.

1.9.3 Supplementary Materials for section 1.4

Stability of Matching and Aggregate Demand Fixing location choice ℓ, the set of feasible schools are defined as $\mathcal{J}_i(\ell;\rho_i) = \{j|c_{ij}(\ell;\rho_i) \geqslant \bar{c}_j\}$, i.e. schools of which student i can clear the cutoffs. The set of feasible schools changes depending on which location ℓ student i chooses to reside in, and the lottery number.

Using the stability of matching and the distributional assumption on the idiosyncratic preference shock over locations and schools,

$$D_j(\{\bar{c}_{j'}\}_{j'=1}^{J}) = \int_i \Sigma_\ell \underbrace{\frac{\exp(\tilde{V}_i(\ell))}{\Sigma_{\ell'}\exp(\tilde{V}_i(\ell'))}}_{\text{Demand for location } \ell} \cdot \underbrace{\int_{\rho_i} \frac{\mathbb{1}(c_{ij}(\ell;\rho_i) \leqslant \bar{c}_j)\exp(\tilde{U}_i(j,\ell))}{\Sigma_{j'}\mathbb{1}(r_{ij'}(\ell;\rho_i) \leqslant \bar{c}_{j'})\exp(\tilde{U}_i(j',\ell))} d\rho_i}_{\text{Demand for school j given location } \ell} \tag{1.18}$$

where

$$\tilde{V}_i(\ell) = \underbrace{W_\ell}_{\text{location char.}} \alpha_i^w + \underbrace{p_\ell}_{\text{housing cost}} \alpha_i^p + \underbrace{\xi_\ell}_{\text{unobs. amenities}} + \alpha^u \underbrace{\mathbb{E}_{\varepsilon_{ij},\rho,\varepsilon_{i\vartheta}}}_{\text{expected util. from school}} \tag{1.19}$$

$$\tilde{U}_i(j,\ell) = \underbrace{X_j}_{\text{school char.}} \beta_i^X + \underbrace{d_{\ell j}\beta_i^d}_{\text{commuting cost}} \tag{1.20}$$

The second component of Equation 1.18 means that a student has effective demand only when school j is feasible given location ℓ. For schools that use lottery number to break the tie, the feasibility depends on the lottery number ρ_i. We take the numerical integration over ρ_i. The existence and the uniqueness of the equilibrium follow from Azevedo and Leshno (2016a). The key assumption is that the distribution of students' priority rank c_{ij} is continuous. This ensures a small change in the cutoff $\bar{c}_{j'}$ induces a small change in the demand for school j.

Stability of Matching and Expected Utility from School In addition, based on the stability of assignment under DA with truth-telling, we can simplify the indirect utility from school choice stage $U_i^*(\ell)$:

$$U_i^*(\ell) = \max\{\max_{j \in \mathcal{J}_i(\ell;\rho_i)} U_i(j_i, \ell; \varepsilon_{ij}), U_i^\vartheta(\vartheta_p; \varepsilon_{i\vartheta}), U_i^\vartheta(\vartheta_c; \varepsilon_{i\vartheta})\}) \tag{1.21}$$

With the assumption that $\varepsilon_{ij}, \varepsilon_{i\vartheta}$ follows i.i.d EVT1 distribution, the expected utility from school can be simplified as follows.

$$\mathbb{E}_{\varepsilon_{ij},\rho,\varepsilon_{i\vartheta}}$$

$$= \mathbb{E}_{\rho}\left(\log\left(\Sigma_{j\in\mathcal{J}_i(\ell;\rho_i)}\exp(\tilde{U}_i(j,\ell;\varepsilon_{ij})) + \exp(\tilde{U}_i^{\vartheta}(\vartheta_p;\varepsilon_{i\vartheta})) + \exp(\tilde{U}_i^{\vartheta}(\vartheta_c;\varepsilon_{i\vartheta}))\right)\right) + \mu \quad (1.22)$$

where

$$\tilde{U}_i(j,\ell) = \underbrace{X_j}_{\text{school char.}}\beta_i^X + \underbrace{d_{\ell j}\beta_i^d}_{\text{commuting cost}} \quad (1.23)$$

$$\tilde{U}_i(\vartheta;\varepsilon_{i\vartheta}) = \beta_i^{\vartheta} \quad (1.24)$$

and μ is the Euler scalar.

Estimating Program Preferences s_{ij} is a weighted average of students' middle school GPA, NYS math and ELA score, and punctuality record, with the weights remaining as each program's private information.

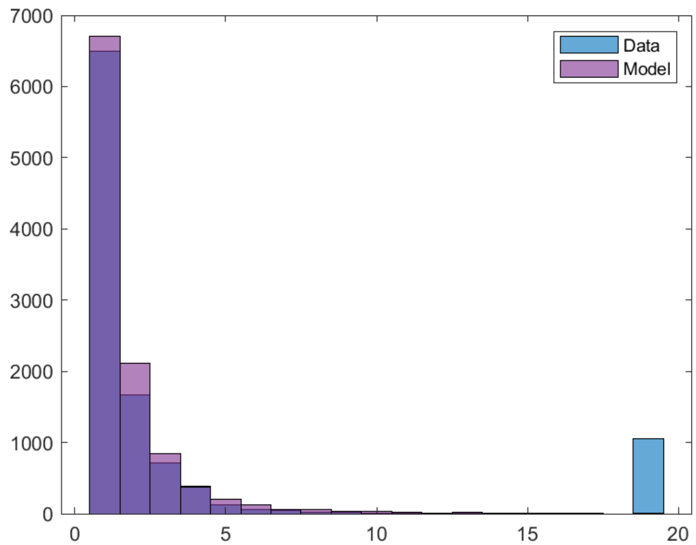
However, a program j has to report the rank of s_{ij}s among its applicants for NYC DOE to implement the centralized DA. Therefore, given any pair of students i and i', we observe the value of $1(s_{ij} > s_{i'j})$, if both i and i' apply to the program j. Using this, we construct $\hat{s}_{ij}$ using a latent variable model by assuming

$$s_{ij} = Z_i\kappa_j + \eta_{ij},\ \eta_{ij} \sim \mathcal{N}(0,1) \quad (1.25)$$

where Z_i is student characteristics that are known to compose of s_{ij}, and κ_j a vector of weights which vary across js. η_{ij} is normalized to be $\mathcal{N}(0,1)$.

We estimate $(\kappa_j)_j$ using Maximum Likelihood Estimation (MLE) where the likelihood function is

$$\mathcal{L}\mathcal{L}_s = \Pi_{i,i'\in\mathcal{A}_j}\Pr(s_{ij} > s_{i'j})^{1(s_{ij}>s_{i'j})}(1-\Pr(s_{ij} > s_{i'j}))^{1(s_{ij}<s_{i'j})} \quad (1.26)$$



Note: Rank is the preference rank of the assigned program. Rank 19 in the data means that the student did not get any offer from programs included in the submitted list.

Figure 1.14: Rank of Assigned School: Model and Simulation

where $\mathcal{A}_j$ is the set of applicants to program j that uses a non-random tie-breaker.

Figure 1.14 shows that our simulation recovers the distribution of the preference rank of the assigned program in the data. Both in simulation and data, around 63% of students are assigned to the 1st- or the 2nd- ranked programs while 8% do not get any offer from programs on the list.

1.9.4 Supplementary Materials for section 1.5

Full Information Maximum Likelihood Estimation Assuming idiosyncratic preference shocks over locations, schools, and outside options ($\eta_{i\ell}, \varepsilon_{ij}$, and $\varepsilon_{i\vartheta}$) are i.i.d, the full information log-likelihood function is as follows.

$$\mathcal{LL} = \underbrace{\Sigma_i}_{\text{3. sum over } i} \log\Big(\underbrace{\Sigma_{k=1}^{K} q_k}_{\text{2. sum over type}} \underbrace{P^{LC}(x_i; \theta^{EC}, \theta^{SC}, \theta^{LC}, \gamma_k) P^{SC}(x_i; \theta^{SC}, \gamma_k) P^{EC}(x_i; \theta^{SC}, \theta^{EC}, \gamma_k)}_{\text{1. likelihood with fixed type}} \Big) \quad (1.27)$$

x_i is data, θ^{LC} is the set of location preference parameters, θ^{SC} is the set of school preference parameters, θ^{EC} is the set of outside option preference parameters, γ_i is the unobserved taste, and q_k the probability of each type k. $(\theta = (\theta^{LC}, \theta^{SC}, \theta^{EC}), \{\gamma_k, q_k\}_k)$ is the full set of parameters to be estimated.

The likelihood function for each step is not additive separable because of γ_i, making the maximization problem computationally very costly. Note that the sequential estimation strategy in Rust (1994) is also not applicable without additive separability of likelihoods.

EM Algorithm with Sequential Maximization The conditional probability of type k given data x_i and current guesses of parameters, are derived using Bayes rule.

$$q(k|x_i; \hat{q}, \hat{\gamma}, \hat{\theta}) = \frac{\hat{q}_k q(x_i; k, \hat{q}, \hat{\gamma}, \hat{\theta})}{\Sigma_{k'} \hat{q}_{k'} q(x_i; k', \hat{q}, \hat{\gamma}, \hat{\theta})} \quad (1.28)$$

We estimate the model using the following iterative process.

1. Initial guess of q^0, γ^0, θ^0

2. Calculate conditional probability in Equation 1.28 using the initial guess q^0, γ^0, θ^0

3. Solve maximization problem for q. It has a closed-form solution which is

$$q_k^1 = \frac{1}{I} \Sigma_i q(k|x_i; q^0, \gamma^0, \theta^0) \quad (1.29)$$

4. Taking other parameters as given, solve the maximization problem of Equation 1.12. Get θ^1_{LC} using the Generalized Method of Moments (GMM) procedure.

5. Taking other parameters as given, solve maximization problem of Equation 1.13 using MLE, get θ^1_{SC} and γ^1.

6. Taking other parameters as given, solve the maximization problem of Equation 1.14. Get θ^1_{EC} using MLE.

7. Repeat 2-6 until convergence

Likelihood Function We presents the likelihood function P^{LC}, P^{SC}, and P^{EC} in this section.

For convenience, we introduced some notations.

$$\tilde{u}_{i\vartheta} = Z_i \beta_\vartheta \tag{1.30}$$

$$\tilde{u}_{ij} = X_j \beta_{Xi} + c(d_{\ell_i j}, Z_i) \tag{1.31}$$

$$\tilde{v}_{i\ell} = W_\ell \alpha_{wi} + p_\ell \alpha_{pi} + \xi_\ell + \mathbb{E}_{\varepsilon_{ij}, \rho_i, \varepsilon_{i\vartheta}} \{U_i^*(\ell) | \ell\} \tag{1.32}$$

Each denotes utility from outside option ϑ, school j, and location ℓ at each decision stage net of idiosyncratic preference shocks.

With the distributional assumption on $\eta_{i\ell}$, the likelihood of location choice P^{LC} takes a simple form.

$$P^{LC} = \Pi_i \frac{\exp(\tilde{v}_{i\ell_i})}{\Sigma_{\ell'} \exp(\tilde{v}_{i\ell'})} \tag{1.33}$$

where ℓ_i is the observed location choice of household i.

We can extend the formula to construct the likelihood of school application P^{SC}. Assuming logit shock and weak truth-telling among eligible options,

$$P^{SC} = \Pi_i \underbrace{\frac{\exp(\tilde{u}_{ij_1^i})}{\Sigma_{j' \in \tilde{\mathcal{J}}_i} \exp(\tilde{u}_{i\ell'})}}_{\text{first-ranked}} \underbrace{\frac{\exp(\tilde{u}_{ij_2^i})}{\Sigma_{j' \in \tilde{\mathcal{J}}_i / j_1^i} \exp(\tilde{u}_{i\ell'})}}_{\text{second-ranked}} \cdots \underbrace{\frac{\exp(\tilde{u}_{ij_{l_i}^i})}{\Sigma_{j' \in \tilde{\mathcal{J}}_i / \{j_1^i, j_2^i \cdots j_{l_i - 1}^i\}} \exp(\tilde{u}_{i\ell'})}}_{l_i^{th}\text{-ranked}} \quad (1.34)$$

where i's observed ranked-ordered list is $\{j_1^i, j_2^i, \cdots, j_{l_i}^i\}$, and $\tilde{\mathcal{J}}_i$ is the set of eligible options for i. Each term in the product represents the probability of choosing the option among eligible options that are not ranked higher.

Similarly, the likelihood of enrollment decision is a product of each i's likelihood, which takes different forms depending on the observed enrollment choice.

$$P^{EC} = \Pi_i \underbrace{\left(\frac{\Sigma_{j' \in \mathcal{J}_i} \exp(\tilde{u}_{ij'})}{\exp(\tilde{u}_{i\mu_i})}\right)^{\mathbb{1}(\vartheta_i = \mu_i)}}_{\text{adjustment when choose } \mu_i} \frac{\exp(\tilde{u}_{\vartheta_i})}{\exp(\tilde{u}_p) + \exp(\tilde{u}_c) + \Sigma_{j' \in \mathcal{J}_i} \exp(\tilde{u}_{ij'})} \quad (1.35)$$

The second term denotes the probability of choosing the enrollment option out of all available options including outside options. The first component adjusts that the distribution of $\varepsilon_{i\mu_i}$ conditional on being assigned to μ_i should be different from the marginal distribution of ε_{ij} and $\varepsilon_{i\vartheta}$. It is more realistic to have $\varepsilon_{i\mu_i}$ preserved the same in the application and enrollment decision stage, rather than assuming that a new $\varepsilon_{i\mu_i}$ is drawn from the Gumbel distribution with a location of zero and scale normalized to one. Without the first term, the value of private options should be underestimated. The proof is available upon request.

Supplementary Tables

	Main	Additional Effects				
	Type1	Type2	Type3	Black/Hisp	FRL	High-achieving
Panel A: School Demand						
Mean test score	0.139	0.161	0.382	0.165	-	-0.245
Frac. Black or Hisp.	-2.002	-0.430	0.794	2.189	0.126	0.355
Frac. FRL	-0.330	-1.077	0.947	-0.590	0.860	-0.278
Non-safety	-0.036	-0.020	-0.007	0.010	-	-
Distance (mi.)	-0.190	-0.937	-5.331	0.096	0.012	0.029
Prob.	0.294	0.629	0.077	-	-	-
Panel B: Outside Option						
Non-public	-2.088	-	-	0.227	-0.879	0.505
Public Charter	-3.307	-	-	1.685	-0.373	-0.178

Note: Columns are for students' heterogeneity and rows are for school characteristics. FRL stands for Free-or-reduced Lunch eligibility. The fastest driving distance to school is calculated using the Open Route Service (ORS). School non-safety measure is constructed by running a Principal Component Analysis (PCA) on crime incidence of different categories at each school building.

Table 1.17: Demand Estimates from *OnlySC* Version

	(1)	(2)
	OLS	2SLS
Dep. var: Mean utility δ_ℓ		
Housing characteristics	Yes	Yes
Land use	Yes	Yes
Neighborhood characteristics	Yes	Yes
log(UnitHousingPrice)	-0.014	-2.441**
	(0.056)	(0.561)
N	1690	1690
First Stage F-stat		17.76
R2	0.641	0.233
ymean	-0.879	-0.879

Note: Instrument variable is the percent park area and the percent residential area of locations that are 2 miles away but within 3 miles from each location.

Table 1.18: IV Regression for Housing Cost

	1st Component	Unexplained variance (percent)
Eigenvalue	3.246	
Total variance explained	64.920	
Eigenvectors:		
Major Crime	0.423	41.900
Other Crime	0.502	18.280
Non-Crime Incidents	0.413	44.520
Property Crime	0.444	36.110
Violent Crime	0.449	34.570

Note: Source - School Safety Report collected by the New York City Police Department which reports the number of crime cases at each school building. Major crimes include burglary, grand larceny, murder, rape, robbery, and felony assault. Other crimes include many crimes that range in severity such as arson, sale of marijuana, or sex offenses. Non-criminal incidents include actions that are not crimes but disruptive such as disorderly conduct, loitering, and possession of marijuana.

Table 1.19: Principal Component of the School Safety Indices

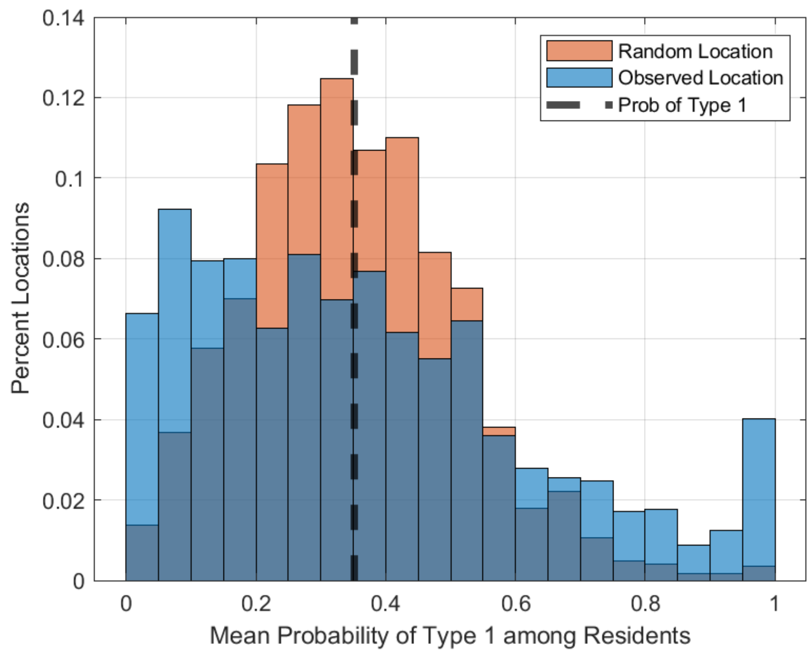


Figure 1.15: Location Sorting on Unobserved Type

	All		Minority		FRL		High-achieving	
	data	model	data	model	data	model	data	model
Panel A1: Mean of Top 3 Ranked School Characteristics								
Mean test score	0.570	0.564	0.215	0.222	0.401	0.404	0.919	0.881
% Black or Hispanic	-0.107	-0.111	0.026	0.024	-0.054	-0.056	-0.212	-0.212
% FRL	-0.061	-0.054	-0.004	-0.001	-0.020	-0.018	-0.115	-0.104
Non-safety	-1.165	-1.091	-0.325	-0.183	-0.851	-0.810	-1.910	-1.766
Distance	1.965	2.032	1.975	2.052	1.949	1.976	1.974	2.065
Panel A2: Mean of Assigned School Characteristics								
Mean test score	0.311	0.352	-0.068	0.016	0.107	0.181	0.704	0.699
% Black or Hispanic	-0.098	-0.087	0.049	0.054	-0.039	-0.029	-0.204	-0.190
% FRL	-0.039	-0.040	0.023	0.017	0.006	-0.001	-0.097	-0.092
Non-safety	-1.037	-0.800	-0.095	0.174	-0.644	-0.493	-1.870	-1.546
Distance	1.382	1.867	1.388	1.940	1.352	1.834	1.455	1.878
Panel B: % Choosing an Outside Option								
Private	5.020	5.844	4.439	5.104	4.187	4.696	5.359	6.359
Charter	4.310	4.903	6.259	7.091	5.079	5.704	3.235	3.724
Panel C: Mean of Chosen Location Characteristics								
log(UnitPrice)	12.846	12.847	12.702	12.702	12.772	12.772	12.963	12.965
% Black or Hispanic	0.568	0.568	0.731	0.731	0.631	0.631	0.459	0.458
Median HH Income	10.696	10.696	10.562	10.561	10.604	10.603	10.814	10.814
Med. travel time to Work	0.759	0.759	0.772	0.772	0.770	0.770	0.747	0.747

Table 1.20: Model Fit

1.9.5 Supplementary Materials for section 1.7

Segregation Measure: Theil's H Index Theil's H Index is also known as the Information Theory Index or the Multigroup Entropy Index. In this paper, we closely follow the definition used by the United States Census Bureau to describe housing patterns (Iceland, 2004).[40]

First, the entropy score of the entire economy is calculated as:

$$E = \sum_{r=1}^{R} (\Pi_r) \log(1/\Pi_r)$$

where Π_r is a particular racial group r's proportion in the whole population in the

[40]See https://www.census.gov/topics/housing/housing-patterns/about/multi-group-entropy-index.html

economy. The entropy score measures the diversity in the economy, where higher number indicates higher diversity.

Next, for each school $j = 1, 2, \cdots, J$, the entropy score of j is calculated similarly:

$$E_j = \sum_{r=1}^{R} (\Pi_{r,j}) \log(1/\Pi_{r,j})$$

where $\Pi_{r,j}$ is a racial group r's proportion in the whole population in school j.

Finally, Theil's H index is calculated as the weighted average of deviation of each j's entropy from the entropy score of the entire economy, where the weight is the number of students at each school:

$$H = \sum_{j=1}^{J} \left[\frac{t_j (E - E_j)}{E \cdot T} \right]$$

where t_j is the total number of students in school j, and $T = \sum_{j=1}^{J} t_j$ is the total number of students in the economy. By construction, H is between 0 and 1 where 0 means maximum integration (i.e., all schools have the same racial composition as the whole economy), and 1 means maximum segregation.

Supplementary Figures

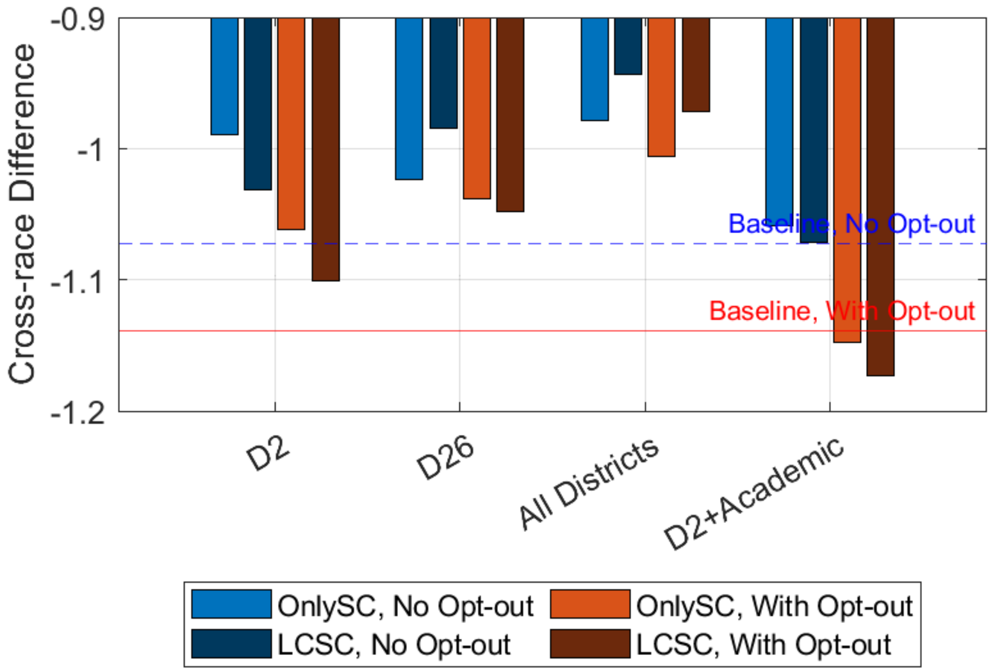


Figure 1.16: Cross-racial Gap in School Characteristics

Note: The figure shows the difference in mean test scores of coassigned/coenrolled peers between Black/Hispanic and other students across scenarios. There are 4 different counterfactual policies. Under **D2**, we introduce purely-lottery based admissions to schools in school District 2. Under **D26**, we target District 26. Under **All District**, we target all schools within the system. Under **D2+Academic**, we scrap location-based admissions rules among District 2 schools while keeping academic screening in place. The dotted line presents the cross-racial gap in coassigned peers' test scores under the status quo. The solid line presents the cross-racial gap in coenrolled peers' test scores under the status quo.

2 A DYNAMIC FRAMEWORK OF SCHOOL CHOICE: EFFECTS OF MIDDLE SCHOOLS ON HIGH SCHOOL CHOICE

2.1 Introduction

Worldwide, numerous jurisdictions employ centralized school choice to assign students to schools at various levels of the education system.[1] The fact that a child faces school choice multiple times makes the interrelated nature of those choices apparent. Anecdotal evidence often suggests that parents are well-aware that a carefully chosen middle school could lead their children to good high schools and consequently to good colleges. In addition, a student's experience at one level (e.g., middle school) may influence which schools she chooses to apply to in subsequent levels (e.g., high school). For example, a student who attends a middle school with high-performing peers may aspire to a high-quality high school placement.

Nevertheless, existing studies rarely examine the dynamic linkages of school choices and often consider each choice in isolation from other choices. If a student's previously chosen schools influence a given school choice, then any analysis that fails to consider these relationships may be contaminated by omitted variables and generate misleading policy implications. For instance, many existing policies that seek to desegregate public high schools focus on reforming only high school admissions. However, segregation patterns in high schools may develop earlier, and such high school-only policies may be insufficient to fully address racial segregation without considering how students' previous schools can influence their high school choices. For instance, if middle schools are already segregated, minority students' middle school experience may differ from that of White students, resulting in different high school application patterns that lead to racial segregation across high schools. Any high school-only admissions reform may fail to address this type of issue.

[1] For example, New York City (NYC) uses the Student-Proposing Deferred Acceptance (SPDA) mechanism (Gale and Shapley, 1962) to assign students to public schools (See Abdulkadiroğlu et al., 2005, 2009, 2017a).

This paper explores the dynamic relationship of school choices at different educational stages and how it affects racial segregation across schools. To our knowledge, this is the first attempt to evaluate the relationship between students' multiple school choices. We ask three questions:

1. Does a student's previous school choice affect the subsequent school choices?
2. If so, what channels mediate this effect?
3. How can one address racial segregation across schools using this relationship?

We use New York City (NYC) public middle and high school choice data to answer these questions. NYC is one of the largest public school districts in the United States, utilizes centralized school choice at multiple levels of education and has been at the center of attention regarding racial segregation across public schools, providing a suitable setting to study the questions.

The analysis proceeds in three parts. First, we start by providing empirical evidence of the causal effects of middle schools on high school applications and assignments. To do so, we use the randomness in the middle school assignments generated by the tie-breaking rules of the admissions system. Second, based on the empirical evidence, we develop and estimate a novel dynamic framework of school choice. We adapt the dynamic discrete choice (DDC) model and combine with large market matching theory to model the dynamics of school choices. Lastly, using the estimated model, we evaluate the impacts of the concurrent admissions reforms on segregation across NYC public schools, when they are implemented at alternative educational stages. In particular, we ask if middle school admissions reform can desegregate not only middle schools but also high schools; this works through the effects of middle schools on high school choice. By doing so, we provide a new perspective on how we should understand and address racial segregation in large school districts.

To elaborate, the first part of the paper provides evidence on the causal effects of the middle schools that students attend on their high school applications and assignments. To overcome students' selection into middle schools based on unobservables, we adopt the research design in Abdulkadiroğlu et al. (2021). The design utilizes the quasi-random assignments to middle schools generated by the tie-breaking rules that distinguish among applicants who have the same applications and same priorities. Our two-stage least squares (2SLS) estimates reveal that all else equal, students who attend high-achievement middle schools *apply to* high school programs that have a higher graduation rate (1.4 pp, 0.08σ), higher college enrollment rate (1.8 pp, 0.10σ), and a higher fraction of high performers (3.0 pp, 0.15σ). Furthermore, such students are *assigned to* programs that have an even higher graduation rate (2.4 pp, 0.13σ), higher college enrollment rate (3.4 pp, 0.18σ), and a higher fraction of high performers (5.3 pp, 0.27σ). Meanwhile, there is no significant effect of attending a middle school with a high fraction of minority students on any outcome.

The second part of the paper turns to a dynamic model of school choice. A model is useful for fully understanding how middle schools shape high school choices for two reasons. First, a student's school assignment is necessarily an equilibrium outcome determined by how *all* students act. While the effects we identified are the marginal effects for each *treated* student, any policy change will trigger a change in the behavior of *all* students. Hence, a model is necessary to analyze the change in the equilibrium induced by any counterfactual policy. Second, having identified the effects of middle schools on high school choice, we are also interested in exploring *how* these effects occur. A model is helpful for decomposing the channels through which middle schools affect high school choice and quantifying each channel's relative importance.

Our two-period model has three key features. First, it explicitly allows how students rank high schools to depend on the middle school they attend (the *application* channel). Second, how high schools rank students for admissions also depends on the middle school that the students attend (the *priority* channel)—for example, middle schools may affect the test scores that students use when they apply to high schools. Third, students are forward-looking; they consider these potential effects on their eventual high school assignments when applying to middle schools. For tractability, we estimate the model using data from Staten Island, which is geographically separated from the other boroughs of NYC. The majority of students do not commute outside of Staten Island, and it can thus be considered an independent school district.

The model estimates have three main implications. First, we reconfirm that the middle school a student attends affects her high school application. For example, all else equal, students who attend high-achievement middle schools are willing to travel an additional 0.11 miles for a 10 pp increase in the proportion of high performers in a high school program. Also, attending a middle school with many students of the same race makes students value the proportion of the same race students in high schools even more. Second, students exhibit forward-looking behavior. That is, students value middle schools that enable them to enjoy a higher expected payoff in high school admissions. Our goodness-of-fit measures show that the model fits the data better than a simple static model in which students are myopic. Third, the unobserved tastes on middle and high school characteristics are serially correlated. The estimates imply that from 18.5 to 28.6% of the variance of the unobservable tastes on high school characteristics can be attributed to the unobservable tastes on middle school characteristics. It suggests that students' selection based on unobservable tastes across two periods exists.

We next turn to investigate how the effects of middle schools on high school choice occur. A decomposition exercise shows that the *application* channel is quantitatively more important than the *priority* channel. That is, the change in high school assignments induced by the change in how students rank high schools is larger than that induced by the change in how students are ranked by the high schools in the admission process. For example, on average, counterfactually changing a student's middle school assignment from the lowest- to the highest-achievement middle school increases the fraction of high performers at her high school assignment by 9.7 pp, and the *application* channel alone explains two-thirds of the increase.

Our findings have both bad and good news about segregation. First, the bad news is that middle school segregation has a reinforcing effect on high school segregation. In terms of academic segregation, we find that students who attended *high quality* middle schools aspire to *higher quality* high schools. In terms of racial segregation, we find that attending middle schools with many students of the same race strengthens racial homophily. These taste changes do impact students' high school assignments through the change in their high school applications. Second, however, the good news is that because students' high school applications and assignments are affected by the middle schools they attend, we may be able to address high school segregation by desegregating middle schools.

To this end, we evaluate the effects of affirmative action policies with different timings in which we eliminate the schools' selection of students so that students' applications and lottery tie-breakers fully decide admissions.[2] We evaluate and compare the impacts of three alternative interventions with different timings on segregation: first, we reform only middle school admissions; second, we reform only high school admissions; and finally, we reform both middle and high school admissions. Our main finding is that the middle school-only reform can desegregate middle schools effectively, and this alters students' applications and assignments to high schools and hence can desegregate high schools as well. For instance, in terms of the gap in the quality of assigned schools between Black/Hispanic and White/Asian students, the middle school-only reform can reduce the gap by 40% in middle schools and at the same time by 10% in high schools. Furthermore, combining both middle and high school reforms has a larger effect on desegregating high schools than reforming only high school admissions.

A more general implication is as follows. Most existing policies that seek to address segregation have focused on reforming the *supply side*, i.e., how schools select students.[3] However, much less attention has been given to how one might address segregation by influencing the *demand side*, i.e., which schools students apply to. Changing only the supply side might not be enough to change students' school assignments and hence address overall segregation. Intervening to change students' applications might have received less attention because, to date, we

[2]This form of policy is equivalent to combining the two recent affirmative action policies the NYC DOE announced in the academic year 2020-21: first, removing screening based on test scores, and second, removing geographic priority rules.

[3]For example, Chicago exam schools (Ellison and Pathak, 2021) use an affirmative action policy that prioritizes students based on the socioeconomic status of their where they reside. Recently, Boston exam schools also adopted a similar admission policy reform (Barry, Ellen, "Boston Overhauls Admissions to Exclusive Exam Schools", The New York Times, 15 July 2021).

have not known of an effective way to influence the demand side.[4] Our findings imply that *early* intervention on the supply side can alter demand side behavior in later periods. We suggest that policy intervention to desegregate high schools should start early and that reforming middle school admissions may be one way to implement such an intervention.

The paper is primarily related to three strands of the literature. First, we add to the economics of education and labor economics literature on the effects of schools on students' future outcomes. Many researchers have studied the effects on outcomes such as academic performance, including test scores (Hastings and Weinstein, 2008b; Jackson, 2010; Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu et al., 2014) or graduation and college outcomes (Deming et al., 2014; Dobbie and Fryer, 2014), or labor market outcomes such as occupation or wages (Card and Krueger, 1992b,a; Betts, 1995; Hoekstra, 2009; Clark and Bono, 2016), among many others. To the best of our knowledge, we are among the first to evaluate the effects of schools on students' future academic choices in a K-12 context.[5] Given the importance of schools for future outcomes as past studies have found, it is crucial to understand what may impact the school attended itself, and we suggest that a student's previous schools may be one key factor.

Second, we contribute to the school choice literature. Several papers have studied the factors that may influence the outcomes of school choice, such as the assignment mechanism (Abdulkadiroğlu et al., 2015b; Abdulkadiroğlu et al., 2017a; He, 2017; Agarwal and Somaini, 2018b; Che and Tercieux, 2019b; Calsamiglia et al., 2020) or information provision (Hastings and Weinstein, 2008b; Hoxby and Turner, 2015b; Ajayi et al., 2017; Luflade, 2017; Corcoran et al., 2018a; Chen and He, 2021a,c; Grenet et al., 2021). However, all these papers were in a static framework and

[4]Few exceptions include providing information to students to make better choices. See the related literature below.

[5]Recently, Mark et al. (2021) conducted a descriptive analysis the within-school and neighborhood similarity in high school applications in NYC, finding low similarities. The key difference is that they treat school programs as distinct objects, while we treat them as bundles of characteristics. We argue that attending different middle schools has a systemic effect on how students view those characteristics, while the exact identities of high schools they apply to may differ. Furthermore, we make use of a quasi-random experiment and a structural model to provide causal evidence.

to our knowledge, we are the first to incorporate a dynamic framework into the school choice literature. To do so, we adopt the dynamic discrete choice (DDC) framework, which in the economics of education literature has primarily been used to study decentralized choice situations such as college major choices (see Arcidiacono and Ellickson, 2011; Altonji et al., 2016, for surveys). We adapt this methodology and combine it with large market matching theory to apply it to the context of centralized school choice. Using this novel dynamic framework of school choice, we add to the literature by explicitly studying the dynamic relationships between school choices made at different educational stages.

Third, we relate to the literature that leverages the quasi-experimental features built in student assignments, which includes making use of lotteries in charter school admissions (Hoxby and Rockoff, 2004), the tie-breaking features of centralized assignments (Deming et al., 2014; Abdulkadiroğlu et al., 2017b, 2021), and the use of test scores generating cutoffs (Hoekstra, 2009; Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu et al., 2014; Dobbie and Fryer, 2014), among many others. We adopt the methodology of Abdulkadiroğlu et al. (2017b, 2021) to obtain the 2SLS estimates of middle schools' causal effect on high school choice. We use students' high school application patterns and assignment results as the outcome variables, departing from Abdulkadiroğlu et al. (2017b, 2021) who study the effect of schools on students achievement.

The rest of the paper is organized as follows. Section 2.2 provides the institutional background for NYC public school choice and describes the data that we use. Section 2.3 analyzes the causal effects of middle schools on high school choice, and Section 2.4 describes our structural model and provides the results of its estimation and the counterfactual analysis. Finally, Section 2.5 concludes and discusses future work.

2.2 Institutional Background and Data

2.2.1 Public School Choice in NYC

NYC is one of the largest school districts worldwide that utilizes centralized school choice to assign students to public schools. The school choice starts as early as 3-K, and students/parents may express their choices in subsequent levels, including Pre-K, kindergarten, elementary, middle and high schools as long as they reside in NYC and wish to enroll in public schools. Schools that are part of the centralized choice system are governed and funded by the city Department of Education (DOE).

This paper focuses on middle and high school choices in NYC. The NYC public middle school system consists of nearly 700 programs at around 500 middle schools. Multiple programs may be offered in one middle school.[6] Middle schools are classified into three types—district schools, borough schools and citywide schools—and a student's residence or elementary school decide eligibility at each type of school.[7] Middle school programs can be further classified into subgroups depending on the admission method.[8] Next, the NYC public high school system consists of nearly 800 programs at around 400 high schools[9]. By contrast, the school choice in high schools is fully citywide—students are eligible at almost all high school programs in NYC. There are multiple admission methods that high school programs use.[10]

Both middle and high school choices use the Student Proposing Deferred Acceptance (SPDA or DA) algorithm (Gale and Shapley, 1962; Abdulkadiroğlu et al., 2005, 2009) which takes students' applications, schools' preferences, and the pre-

[6]Multiple programs may be offered in one high school too. Since the unit of admission is program instead of school, one may consider each program as a separate school. In the following, we use the term 'program' and 'school' interchangeably when there is no confusion.

[7]For example in academic year 2014-2015, there were 670 programs in 472 schools. Among them, 14 programs were citywide school programs, 39 programs were borough school programs.

[8]See section 2.6.1 for details on the admission methods of middle and high schools.

[9]In academic year 2017-18, there were 767 programs in 426 schools.

[10]Additionally, there are 9 specialized high schools in NYC, for example, Stuyvesant High School or Bronx High School of Science. We exclude these specialized high schools from our analyses since they use a separate admission process using a test called Specialized High Schools Admissions Test (SHSAT).

announced number of seats as main inputs and produces at most one assignment for each student.[11]

Students apply to programs by subtmitting rank-ordered list (ROL). In middle school choice, students are allowed to rank however many programs at which they are eligible. In high school choice, students may rank up to 12 choices of high school programs.[12]

Next, how schools rank students (schools' preferences) is decided by pre-announced admission rules, consisting of three layers. First, eligibility criteria decide at which programs a student is eligible. Second, eligible applicants are classified into a small number of priority groups, for example, 'students or residents of Manhattan' or 'students who attended the information session'. A student in the higher priority group is always considered before any student in lower priority groups in the admission process. Finally, a tie-breaking rule is often required since the number of priority groups is much smaller than the number of applicants, and multiple applicants thus belong to the same priority group. How the ties are broken depends on the program's admission method. First, any type of 'screening' programs including Screened, Screened: Language, Screened: Language & Academics, Audition, Talent Test, and seats reserved for screening at Educational Option programs breaks tie using a program-specific *non-random tie-breaker*, which usually consists of previous year's GPA, statewide standardized test scores, attendance, and punctuality. Next, other programs that do not screen students break ties by a random *lottery* which is attached to each student and applies to all such programs in the same fashion (single tie-breaking rule).

[11] See subsection 2.6.2 for details on how SPDA works.

[12] In this regard, the algorithm used for high school assignment is a modified version of SPDA with a limit on the number of choices, which alters the nature of SPDA (Haeringer and Klijn, 2009b; Calsamiglia et al., 2010). For example, strategyproofness does not hold. However, we do not rely on the strategyproofness of SPDA throughout this paper.

2.2.2 NYC School Choice Data

We focus on the main round application data of students who participated in the middle school application (MSAP) in the academic year 2014-15, and then participated in high school application (HSAP) in the academic year 2017-18.[13] subsection 2.6.1 provides more details on data sources and sample restrictions. We have 54,012 students applying to 670 middle school programs (472 middle schools) in the academic year 2014-15 and 767 high school programs (426 high schools) in the academic year 2017-18 as our main sample.

In the following analysis, we focus on two types of schools—1) high achievement, and 2) high minority. A school is labeled 'high achievement' if the average standardized test score of current students belongs to the top 1/3 in the distribution across all schools. Similarly, a school is labeled 'highly minority' if the proportion of Black and Hispanic students belongs to the top 1/3 in the distribution across all schools.[14] School types are defined based on the characteristics of the current seniors in the academic year 2014-15 and 2017-18, respectively. These types are neither exclusive nor exhaustive.

We present summary statistics of baseline student characteristics in Table 2.1. Columns (1)-(3) present summary characteristics of all middle school applicants (whole sample), and Columns (4)-(6) present those of middle school applicants after attrition (main sample). The majority of students are either Black (23%) or Hispanic (41%) and Free or Reduced Lunch (FRL) eligible (72%). 53% of students ranked a high achievement middle school as their first choice. On average, 1.7

[13]The timeline of the admission process is as follows (Corcoran and Levin (2011b), Directory of NYC Public High Schools). By December, students are required to submit their ROLs. By March, SPDA algorithms are run which determines students' assignments (the 'main' round). Students who accept their offer finalize, and if a student rejects an offer then she goes to the next round. This describes the main round of the entire system. A majority of students finalize in the main round (about 85% each year). Students who are not assigned in the main round or rejected the assignment go to the Supplementary round which is similarly organized as the main round and includes school-programs that did not fill up their capacities in the main round, or programs that are newly opened. Finally, there is an administrative round in which students who are not assigned a school even after the second round are administratively assigned a school.

[14]We find the main results are not sensitive to a different definition of types, for example, using above median, 60th-, 70th-, and 75th-percentile in the respective distribution.

schools on a student middle school ROL are high achievement middle schools. There is a remarkable variation from one student to another, captured by the sizable standard deviations. Finally, the demographic characteristics and middle school application behavior are very similar between the whole sample and the main sample.

Variables	(1) N	(2) Mean	(3) Std	(4) N	(5) Mean	(6) Std
	All MS applicants			Both MS and HS Application		
5th Grade ELA score	62,972	300.2	35.4	54,012	300.6	35.0
5th Grade Math score	62,972	310.8	37.7	54,012	311.3	37.3
English Language Learner (ELL)	62,972	0.12	0.32	54,012	0.12	0.32
Special Education Status	62,972	0.21	0.41	54,012	0.21	0.40
Free or Reduced Lunch (FRL)	62,972	0.72	0.45	54,012	0.73	0.45
Female	62,972	0.49	0.50	54,012	0.50	0.50
Asian	62,972	0.18	0.39	54,012	0.19	0.39
Black	62,972	0.23	0.42	54,012	0.23	0.42
Hispanic	62,972	0.41	0.49	54,012	0.41	0.49
White	62,972	0.17	0.37	54,012	0.17	0.37
Ranked High Achievement MS 1st?	63,207	0.53	0.50	53,211	0.53	0.50
# of High Achievement MS Ranked	63,207	1.66	1.71	53,211	1.67	1.72
Ranked High Minority MS 1st?	63,207	0.20	0.40	53,211	0.20	0.40
# of High Minority MS Ranked	63,207	0.78	1.46	53,211	0.77	1.44

Notes: Summary statistics of student characteristics in 5th grade are presented. A middle school is 'high achievement' ('high minority') if the average standardized test score (the percent of Black and Hispanic students) of current students is greater than the 66th percentile of that across all schools. The scale of 5th Grade ELA score is from 100 to 410, and that of 5th Grade Math score is from 130 to 420.

Table 2.1: Summary Statistics of Student Characteristics

Table 2.2 shows summary statistics of programs on admissions criteria and enrolled students' characteristics, overall and by school type. While 94% of middle school programs are open only to students from a school district or an attendance zone, only 4% of high school programs employ such priority rules. Next, many middle and high schools employ non-random tie-breakers to cream-skim high

Variables	(1) Mean	(2) Std	(3) Mean	(4) Std	(5) Mean	(6) Std
	All		High Achievement		High Minority	
Panel A: Middle School Program Characteristics						
Open Only to District/Zone Students?	0.94	(0.24)	0.89	(0.32)	0.96	(0.21)
Use Non-random Tie-breaker?	0.42	(0.49)	0.59	(0.49)	0.35	(0.48)
Average Test Score (6th)	297.3	(20.5)	313.0	(17.7)	282.5	(11.8)
% Female	49.5	(12.52)	52.24	(15.34)	48.50	(10.83)
% White	14.17	(20.88)	27.51	(25.27)	1.062	(1.78)
% Black or Hispanic	70.92	(30.51)	47.92	(30.82)	97.16	(2.74)
% Free or Reduced Lunch	76.09	(19.06)	66.70	(22.05)	87.48	(8.75)
Cohort Size	98.30	(90.67)	111.90	(103.00)	71.08	(37.07)
STEM	0.14	(0.35)	0.13	(0.33)	0.17	(0.37)
Panel B: High School Program Characteristics						
Open Only to District/Zone Students?	0.04	(0.19)	0.05	(0.21)	0.01	(0.09)
Use Non-random Tie-breaker?	0.38	(0.49)	0.61	(0.49)	0.31	(0.46)
Graduation Rate (%)	73.19	(15.98)	85.19	(9.88)	68.43	(15.12)
4yr Graduation Rate (%)	67.01	(17.34)	81.79	(10.81)	60.67	(16.19)
College Enrollment Rate (%)	58.38	(17.13)	73.96	(11.49)	52.23	(15.11)
Average Test Score (9th)	294.4	(17.49)	311.3	(14.36)	285.7	(13.37)
% Female	49.01	(20.09)	53.89	(19.22)	48.99	(20.00)
% White	10.41	(15.56)	19.47	(19.69)	1.629	(02.15)
% Black or Hispanic	76.58	(23.40)	57.98	(26.10)	95.49	(04.20)
% Free or Reduced Lunch	80.13	(15.33)	70.75	(18.37)	87.78	(08.55)
Cohort Size	83.04	(82.66)	114.50	(122.70)	65.36	(45.48)
STEM	0.31	(0.46)	0.31	(0.46)	0.28	(0.45)
% From High Achievement MS	29.89	(26.68)	58.99	(24.80)	15.69	(17.40)
% From High Minority MS	32.96	(26.53)	15.62	(21.71)	52.78	(22.82)

Notes: A middle school is ‘high achievement’ (‘high minority’) if the average standardized test score (the percent of Black and Hispanic students) of current students is greater than the 66th percentile of that across all schools. Test score is a mean of ELA (English Language Arts) and math test scores. Educational Option high school programs are not counted as non-random tie-breaker. The scale of average test score if from 110 to 410 (130 to 400) for middle school (high school) programs.

Table 2.2: Summary Statistics of Middle and High School Program Characteristics

achievement students. 59% of high achievement middle schools adopt non-random tie-breakers relative to the average of 42%. The contrast is sharper at the high school level (61% vs. 38%).

Table 2.2 also illustrates that enrolled students' characteristics vary markedly depending on school type, suggesting that students sort into different schools based on their characteristics. The mean average test score among all middle schools is 297.3, while it is 313.0 among high achievement middle schools and 282.5 among high minority middle schools. Similarly, while middle schools have 14% of White students on average, a high achievement middle school has 28%, and a high minority middle school has only 1%. The pattern is similar among high schools.

Importantly, the last two rows of Panel B of Table 2.2 show a correlation between the type of middle school a student graduated from and the type of high school she attends. While on average 30% of students in high schools have graduated from a high achievement middle school, the number is twice as large among high achievement high schools. Similarly, high minority high schools admit more students who graduated from high minority middle schools than an average high school. These patterns suggest two possibilities. First, students may have consistent tastes over middle and high school programs. Second, which middle school a student attends may play an important role in how she applies and is assigned to high schools. We aim to explore these possibilities in the following sections.

Next, Table 2.3, Table 2.4 summarize the averages of school characteristics by rank on students' ROLs of middle schools and high schools, respectively. There are mainly three patterns. First, students tend to rank distant schools from their homes lower on their ROLs. Notably, the average distance of ranked programs is larger for high school programs than for middle school programs. As mentioned above, this possibly reflects that high school application has a higher degree of citywide school choice. Next, students rank schools with high student achievement higher on their ROLs. Third, students rank schools with a high fraction of subsidized lunch status, Black or Hispanic students lower on their ROLs.

	1	2	3	4	5	6	7	8	9	10	11	12 or longer
Number Ranked	52789	40428	33980	24435	13419	7439	4131	2859	2013	1557	961	729
% Ranked	97.7	74.9	62.9	45.2	24.8	13.8	7.6	5.3	3.7	2.9	1.8	1.3
Distance (miles)	1.4	1.7	1.8	1.9	1.9	2	2.1	2.3	2.4	2.6	3	3.2
Mean Score (6th grade)	308.2	307.7	305.8	305.1	305.7	305.8	306.2	306.2	305.3	305.8	306.8	304.2
Mean Score (8th grade)	300.7	300.4	299.2	298.5	299.1	299.6	300.7	300.3	299.3	299.2	299.6	296
% Black or Hispanic	63.8	65.9	69.5	69.2	67.3	67.5	66.5	65.6	65.7	68.2	63.8	70.3
% Female	49.9	50.3	50.1	50.1	49.7	49.9	49.9	49.6	49.4	49.6	50.2	49.9
% Free or Reduced Lunch	69.4	69.9	71.6	72.2	72.7	73.7	74	73.9	75.1	73.8	71.8	73
6th Grade Size (100s)	1.6	1.3	1.2	1.2	1.1	1.1	1.1	1.1	1.1	1	1.1	0.9

Notes: The table calculates the average characteristics of the middle school programs on students' ROLs, by the rank on the ROL (N=54,012). % Black or Hispanic, % Female and % Free or Reduced Lunch are calculated using the characteristics of the currently enrolled 6th graders in AY 2014-15. Mean Score (6th grade) and Mean Score (8th grade) are calculated using the average of the statewide standardized Math and ELA exams of currently enrolled 6th graders and 8th graders in AY 2014-15, where the scale is from 110 to 410.

Table 2.3: Middle School Program Characteristics on ROLs

	1	2	3	4	5	6	7	8	9	10	11	12
Number Ranked	53187	49070	47234	44381	41062	37011	32413	28235	23943	20435	16952	13402
% Ranked	98.5	90.9	87.5	82.2	76.0	68.5	60.0	52.3	44.3	37.8	31.4	24.8
Distance (miles)	3.2	3.4	3.5	3.6	3.7	3.8	3.9	3.9	4.0	4.0	4.0	3.9
Mean Score (9th grade)	312.7	310.8	309.1	307.9	307.1	306.0	305.6	304.7	304.1	303.2	302.5	301.2
4yr Grad Rate (%)	85.4	84.1	83.4	82.8	82.5	82.1	82.0	81.6	81.4	80.9	80.2	79.2
Enroll in College (%)	73.8	72.3	71.3	70.7	70.2	69.7	69.6	69.1	68.8	68.1	67.3	66.1
% Black or Hispanic	58.2	59.5	60.7	62.4	63.5	65.0	66.0	67.4	68.5	69.9	70.7	71.6
% Female	53.4	51.9	51.1	50.7	50.4	50.2	50.1	50.0	49.7	49.8	49.7	49.5
% Free or Reduced Lunch	69.8	71.2	72.2	73.3	73.8	74.5	74.9	75.5	76.0	76.6	77.4	78.0
9th Grade Size (100s)	1.7	1.5	1.5	1.4	1.4	1.3	1.3	1.3	1.3	1.2	1.2	1.3

Notes: The table calculates the average characteristics of the high school programs on students' ROLs, by the rank on the ROL (N=54,012). % Black or Hispanic, % Female and % Free or Reduced Lunch are calculated using the characteristics of the currently enrolled 9th graders in AY 2017-18. Mean Score (9th grade) are calculated using the average of the 8th grade statewide standardized Math and ELA exams of currently enrolled 9th graders in AY 2017-18, where the scale is from 130 to 400. 4yr Grad Rate and Enroll in College are calculated using the average of the graduating cohort in AY 2017-18.

Table 2.4: High School Program Characteristics on ROLs

2.3 Causal Effects of Middle School Attendance on High School Choice

In this section, we provide evidence on the causal effects of middle schools on high school choice. Specifically, we show that which middle school a student attends plays an important role in the subsequent high school application behavior and the assignment results. To this end, we estimate the treatment effects of attending a given type of middle schools[15] on the characteristics of high school programs a student ranks and is assigned to. The main identification concern is that students may sort into different middle schools through the middle school choice, based on some factors unobservable to the researcher. These unobservable factors might at the same time affect how students make high school choices and where they are assigned to. Without considering the sorting based on unobservables, any estimate on the treatment effects of attending some type of middle schools will be biased. To deal with this selection issue, we adopt the research design established by Abdulkadiroğlu et al. (2017b, 2021). The design builds on the quasi-experimental variation embedded in a centralized SPDA. We explain the strategy briefly in the following, and we recommend that interested readers consult the original papers for more details.

2.3.1 Empirical Strategy

In NYC, public school assignments are solely determined by students' applications, priorities,[16] and tie-breakers. Recall that priorities are often too coarse to decide which students to admit/reject, and programs use two types of tie-breakers in such cases: *lotteries* and program-specific *non-random tie-breaker* (see section 2.2).

First, for a program which uses lotteries, each student is assigned a random lottery number, and the lottery breaks ties when there are more applicants than the

[15]We use two alternative treatment variables: attending a 'high achievement' middle school, and a 'high minority' middle school. See section 2.2 for the definition of those types.

[16]For convenience, we use 'priority' to denote both eligibility and priority groups.

number of seats. Hence after controlling for student application and priority, students' assignments at those programs are random (Abdulkadiroğlu et al., 2017b). On the contrary, for any program that uses a non-random tie-breaker, assignment is no longer random even after controlling for student application and priority since the non-random tie-breaker itself might be correlated with a student's type, for example, unobserved abilities. Abdulkadiroğlu et al. (2021) take a non-parametric regression discontinuity (RD) approach (Hahn et al., 2001) and show that applicants whose priority scores[17] are in the small neighborhood around the cutoffs of such programs have a constant risk of clearing the cutoffs of 1/2 (Proposition 1 of Abdulkadiroğlu et al., 2021), and hence those assignments are as good as random.

However, in practice, it is impossible to control for all observed cases of student application and priority as there are as many unique combinations of applications and priorities as the number of students. Abdulkadiroğlu et al. (2021) show that conditioning on the propensity score—the probability of being assigned to the treatment schools—DA generated assignments are independent of any variables that are unaffected by treatment, and hence eliminating any omitted variable biases and at the same time reducing the dimension effectively (Rosenbaum and Rubin, 1983).[18] Theorem 1 of Abdulkadiroğlu et al. (2021) provides a compact characterization of such propensity scores using a large market approximation. subsection 2.6.3 provides a simple example of the calculation of the propensity scores.

We estimate a two-stage least squares (2SLS) model where the DA *assignment* is used as an instrumental variable for the actual *attendance*, since the actual *attendance*

[17]Priority score is a combination of priority group and the tie-breaker that summarizes each applicant's priority at the program for admissions.

[18]Propensity score denotes the odds of being assigned to a certain type of middle school as a function of student application, priority group, and cutoffs. We can calculate the propensity score for each middle school program for each student. Since SPDA produces at most one assignment for each student, summing up the propensity score across middle school programs that belong to a certain type gives the propensity score of being assigned to middle schools of such type. If a student does not apply any of middle schools of a certain type, the propensity score is zero.

could be different from the DA *assignment*.[19]

$$Y_i = \alpha_0 + \beta C_i + \sum_x \alpha_1(x) d_i(x) + g(\mathcal{R}_i) + \delta' Z_i + \eta_i \quad (2.1)$$

$$C_i = \tilde{\alpha}_0 + \gamma D_i + \sum_x \alpha_2(x) d_i(x) + h(\mathcal{R}_i) + \tau' Z_i + \nu_i \quad (2.2)$$

Equation 2.1 is the main equation of interest where β is the treatment effect of interest, and Equation 2.2 is the respective first-stage regression. Y_i is our outcome of interest describing student i's high school choice behavior or outcomes, C_i is the treatment variable which equals 1 if i attended one of the treatment middle schools and 0 otherwise. D_i is the DA assignment into treatment middle schools which equals 1 if i was assigned to treatment schools by DA and 0 otherwise. We also include Z_i, the vector of student observable characteristics (ELL, ethnicity, free or reduced lunch (FRL) status, gender, baseline test scores, and borough of residence) when they were 5th graders i.e., before applying to middle schools. $\{d_i(x)\}_x$ provides a saturated nonparametric control for all possible values of propensity score for the DA assignment D_i, and $g(\mathcal{R}_i)$ and $h(\mathcal{R}_i)$ are local linear controls for non-random tie-breakers at each program that uses such tie-breakers.[20] [21]

To interpret β as causal, we argue that the exclusion restriction holds. That is, DA assignments D_i have no effect on outcomes Y_i other than by affecting the actual attendance C_i after controlling for propensity scores and non-random tie-breakers. To support this assumption, we provide balance test results in Figure 2.1 showing that the instrumental variable balances the covariates of the students who

[19]Even when DA *assignment* is as good as random conditional on propensity scores, focusing on the effect of the actual *attendance* can re-introduce the selection issue; students can appeal or transfer in order to attend a school different from their assignment.

[20]There are 104 types of non-random tie-breakers in the data, and we include a local linear function for each one. We also include a set of dummy variables corresponding to each non-random tie-breaker to deal with students who did not apply to a school using that non-random tie-breaker, or students who applied but whose tie-breakers are far from the cutoff following Abdulkadiroğlu et al. (2021). We use the IK bandwidth (Imbens and Kalyanaraman, 2012) separately for each program as suggested by Abdulkadiroğlu et al. (2021).

[21]In principle, controlling for the propensity scores $\{d_i(x)\}$ is enough for exclusion restriction by Theorem 1 of Abdulkadiroğlu et al. (2021). The additional controls for student characteristics and non-random tie-breakers contribute to more precise estimate of the treatment effect β.

are assigned to the treatment middle schools by DA (*offered* students) and those who are not (*non-offered* students).

In Figure 2.1 and the following analysis, we focus on the treatment 'attending a high achievement middle school' for the purpose of illustration. We check the balance of students' test scores, demographic characteristics, and variables that describe middle school application behavior.

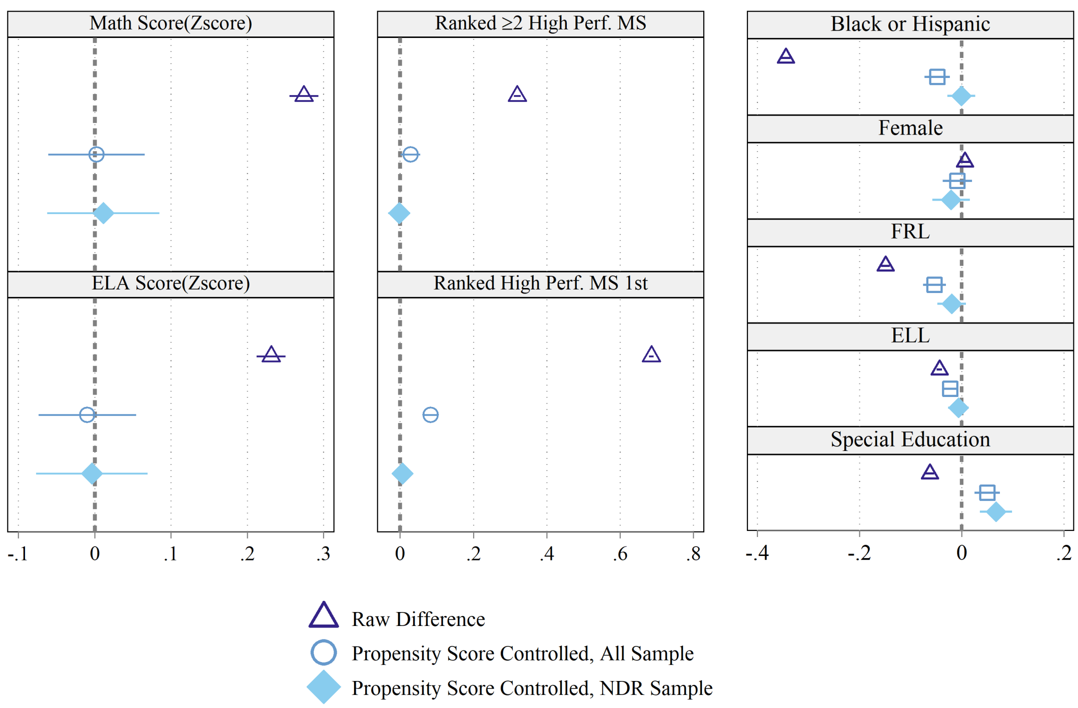
First, **Raw Difference** shows the sharp raw difference of covariates between the offered and the non-offered students. The offered have higher test scores and are less likely to be FRL, ELL, Black or Hispanic, or need special education, with all statistically significant differences. They also rank more high achievement middle schools (recall this is our treatment of interest) than the non-offered students, and are more likely to list them first on their ROLs, which makes natural sense as such behavior will unambiguously increase the odds of being offered such schools.

Next, we control for the propensity scores and include local linear control of tie-breakers in the following two specifications denoted by **Propensity Score Controlled, All Sample** and **Propensity Score Controlled, NDR Sample** in Figure 2.1. Specifically, we run

$$W_i = \alpha_0 + \gamma D_i + \sum_x \alpha_1(x) d_i(x) + h\left(\mathcal{R}_i\right) + e_i \tag{2.3}$$

where W_i is the student covariates that we test balance, and D_i, $\{d_i(x)\}_x$ and $h(\mathcal{R}_i)$ are the same as in our main specification Equation 2.1.

Propensity Score Controlled, All Sample in Figure 2.1 presents estimates on γ with students of all possible propensity scores, including 0 and 1. Controlling for the propensity score and non-random tie-breakers effectively balances covariates. Next, **Propensity Score Controlled, NDR Sample** shows the γ only with students with non-degenerate risk of being offered, i.e. subject to randomization. Further restricting the sample to those with non-degenerate risk provides an almost perfect balance between the offered and the non-offered group. Based on the balance test result, our preferred specification in the following controls for propensity scores and non-random tie-breakers with restricted sample of non-degenerate risk of



Notes: **Raw Difference** shows the t-test results of covariate mean difference between the offered and the non-offered. **Propensity Score Controlled, All Sample** shows the coefficient of the offered when we regress the covariate on the offered dummy and nonparametric control for propensity score and local linear function of non-random tie-breaker using the entire sample. **Propensity Score Controlled, NDR Sample** is similar to **Propensity Score Controlled, All Sample** but when we only include the sample of which propensity score is neither 0 nor 1. The unit is relative difference of each covariate of the offered students to that of the non-offered students, and is standard deviation for the left panel and fraction for the middle and right panels. Markers show the exact estimates, and 95% CIs are presented. Robust standard errors are estimated. N=8,007 for **Propensity Score Controlled, NDR Sample**, and N=50,871 for other estimates.

Figure 2.1: Covariate Balance Test: Offered Students v.s. Non-offered Students

offer.[22]

2.3.2 Empirical Results

Table 2.5 shows our main results. Each panel corresponds to different high school characteristics as the outcome variable.

In Columns (1)-(3), we focus on the average characteristics of the top 5 ranked high school programs.[23] Column (1) presents OLS estimates for comparison. Column (2) presents 2SLS estimates with full sample, and Column (3) presents our preferred specification—2SLS only with non-degenerate risk sample. First, we find the OLS estimates overestimate the effects of attending a high achievement middle school as concerned. For example in Panel C, the OLS estimate suggests that the average fraction of high performers of a student's top 5 ranked high school programs increases by 5.19 percentage points when she attends a high achievement middle school. On the other hand, the 2SLS estimate in Column (2) shows an effect of 3.33 percentage points, and our most preferred estimate in Column (3) is 2.99 percentage points. This contrast confirms the importance of controlling for selection based on unobservables.

Most importantly, our 2SLS estimates illustrate that attending a high achievement middle school has a causal effect on the characteristics of high school programs a student applies to. We see that the average graduation rate, college enrollment rate, and the fraction of high performing students of a student top-5 choice increase by 1.38, 1.76, and 2.99 percentage points, respectively.

Columns (4)-(6) illustrate that attending a high achievement middle school also changes the characteristics of the assigned high school program, not only

[22]Such sample restriction comes with the cost of losing many observations (from N=50,871 to N=8,007). **NDR-DR** in Figure 2.5 presents the mean difference between those with non-degenerate offer risk and degenerate (0 or 1) offer risk. We find that students with non-degenerate risk and those with degenerate risk are quite different: students with non-degenerate risk have higher test scores, and more likely to be White. It reconfirms that the 2SLS estimates we find in the next section are local average treatment effect (LATE).

[23]Using the characteristics of the top choice, or average characteristics of top 3 choices does not significantly change the results.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Average of Top 5			Matched		
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: 4yr Graduation Rate (%)						
From High Achievement MS	1.764***	1.684***	1.379*	3.109***	2.482***	2.422**
	(0.404)	(0.561)	(0.735)	(0.529)	(0.855)	(1.142)
N	44159	44159	7060	41623	41623	6687
R2	0.293	0.318	0.387	0.185	0.202	0.253
$\bar{y}$	83.321	83.321	83.729	78.954	78.954	79.901
Panel B: College Enrollment Rate (%)						
From High Achievement MS	2.854***	1.716**	1.755*	4.530***	3.025**	3.414**
	(0.516)	(0.780)	(1.011)	(0.669)	(1.191)	(1.566)
N	44158	44158	7060	41546	41546	6679
R2	0.367	0.390	0.459	0.244	0.263	0.310
$\bar{y}$	71.217	71.217	72.197	65.653	65.653	67.204
Panel C: % High Performing Students						
From High Achievement MS	5.188***	3.328***	2.986*	6.886***	5.293***	5.292**
	(0.840)	(1.291)	(1.805)	(0.825)	(1.650)	(2.105)
N	44237	44237	7062	42180	42180	6751
R2	0.450	0.473	0.502	0.388	0.406	0.400
$\bar{y}$	39.731	39.731	40.934	33.058	33.058	34.978
Panel D: % White						
From High Achievement MS	5.080***	2.202***	0.311	5.755***	1.915**	0.301
	(0.750)	(0.729)	(0.655)	(0.793)	(0.819)	(0.832)
N	44237	44237	7062	42180	42180	6751
R2	0.633	0.652	0.717	0.555	0.573	0.621
$\bar{y}$	18.627	18.627	20.334	15.097	15.097	16.761
Panel E: 1(STEM)						
From High Achievement MS	-0.053***	0.013	0.041	-0.057***	0.022	0.055
	(0.013)	(0.024)	(0.035)	(0.016)	(0.032)	(0.044)
N	44237	44237	7062	42182	42182	6751
R2	0.098	0.126	0.275	0.041	0.059	0.172
$\bar{y}$	0.324	0.324	0.318	0.314	0.314	0.322
First Stage F-stat		146.8	135.2		146.8	135.2

Notes: Standard errors clustered at graduating middle school in parentheses, *p<0.1, **p<0.05, ***p<0.01. All regressions control for student ethnicity, gender, English Language Learner status, Free or Reduced Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. Column (2)-(3) and (5)-(6) control for saturated dummy for all possible values of propensity score of being assigned to a high achievement MS, and local linear control for non-random tie-breakers.

Table 2.5: Effect of Attending a High Achievement MS on HS Characteristics

of the programs students apply to. First, attending a high achievement middle school changes the quality of the high school program a student is assigned—the graduation rate of the assigned high school program increases by 2.42 percentage points, college enrollment rate by 3.41 percentage points, and the fraction of high performing students by 5.29 percentage points. Second, note that the magnitude of effects is larger with assigned programs than those a student applies to. This implies that attending a high achievement middle school not only changes how students value different program characteristics, but also how a student is viewed by those programs for admission purposes.[24] We build these possible channels of middle schools' effects into our dynamic model in section 2.4.

The previous results confirm the causal effect of middle schools on high school applications and assignments. In section 2.6.8, we find that the effects we identify are robust to controlling for the students' end of middle school test scores and the length of the submitted ROLs. Also, we explore the heterogeneity in treatment effects by student observable characteristics in Figure 2.6. Finally, Appendix Table 2.19 shows the effects of attending a high minority school middle school, which is our second treatment of interest, finding no significant effect.

2.4 A Dynamic Model of Middle and High School Choice

Having confirmed the effects of middle school choice on high school choice, we now develop and estimate a dynamic model of middle school and high school choice.

The need for a model is twofold. First, students' school assignments are determined as an equilibrium outcome in which how *all* students act jointly determines

[24]The figures are slightly larger than Corcoran et al. (2018a) which conducts a field experiment by providing a customized one-page list of proximate high schools with a high graduation rate to high poverty middle schools in NYC. They find the treatment group did not *apply to* high schools of higher quality measured by graduation rate, but was *assigned to* schools with 1.7pp (0.12σ) higher graduation rate.

the outcome. However, the effect we identified in the previous section is marginal for each *treated* student. Once we are interested in analyzing any counterfactual policy change, an equilibrium model is required because it will trigger a change in behavior of *all* students, in turn changing the equilibrium. Second, we are also interested in exploring *how* the effects of middle schools on high school choice we identified occur. A model is useful to decompose the channels through which middle schools affect high school choice and to quantify each channel's relative importance.

The model is a two-period dynamic model. The first period corresponds to middle school applications and assignments, and the second period corresponds to high school applications and assignments. We incorporate four key features in our model.

First, the model explicitly allows students' tastes for high schools that underlie their applications to depend on the middle school they attend (*application* channel).

Second, how a student is prioritized at each high school program for admissions also depends on the middle school she attended (*priority* channel). Test scores can change by attending middle schools with different value-added, and a student's eligibility and priority group at each high school program may also change depending on the middle school she attends.[25]

Third, students are forward-looking; namely, they take the effects on high school choice into consideration when they apply to middle schools. More concretely, students form expectations on how they will benefit in the high school choice by attending a particular middle school, which in turn affects how they value different middle school programs.

Finally, we explicitly take care of selection into middle schools based on unobservables. This shares the same concern as in the previous section. That is, if there are unobservables that affect both middle school and high school choices we fail to control for, any parameter estimate will be biased. We include unobservable tastes on school characteristics for both middle and high schools, and allow them to be

[25]For example, at continuing 8th graders programs which give highest priority to students who are from the same middle schools.

serially correlated to model selection across the two periods.

2.4.1 Theoretical Framework: A Two-Period Model

Behavioral Assumption We need an assumption on the equilibrium behavior of students to interpret the school choice data we observe. We assume that students submit ROLs such that the resulting assignment outcomes are ex-post stable (Che et al., 2021). That is, the assigned program of a student is her favorite program among those that were feasible to the student. This behavioral assumption is consistent with the implication of the truth-telling assumption that has been traditionally used in the school choice literature (for example, Abdulkadiroğlu et al., 2017a) which builds on the strategyproofness of SPDA.[26] Furthermore, it is also consistent with more recent literature on students' departure from weakly dominant strategies even in a strategyproof environment (Hassidim et al., 2016; Li, 2017; Artemov et al., 2021).

Ex-post stability plays a significant role in simplifying a rather complicated game situation. In particular, we can focus on outcomes rather than strategies. That is, it enables us to interpret the school choice data such that for each student, her assigned program gives the maximum utility among the programs that were feasible to the student, without relying on knowing the exact strategy the student employed. Without ex-post stability, one needs to fully solve the game of incomplete information that each student is facing by enumerating all possible ROLs and finding the optimal strategy profile among them, which would make the estimation of the model extremely heavy in terms of computation. Especially, it helps us simplify the continuation value (known as the 'Emax' term in the dynamic discrete choice literature), as will be seen in the description of our model.

In the following, denote each player (student) as $i \in \{1, \cdots, I\} = \mathcal{I}$, middle school programs as $m \in \{1, \cdots, J_m\} = \mathcal{M}$, and high school programs as $j \in \{1, \cdots, J_h\} = \mathcal{J}$. We start from period 2 and work backwards.

[26] That is, for each student, it is a weakly dominant strategy to report her true preferences.

2.4.1.1 Period 2: High School Application

Each forward-looking player (student) i plays a Bayesian game in each of the two periods. In the second period (high school application), students submit ROLs on high schools programs satisfying ex-post stability, based on the flow utilities

$$V_{ij} = v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) + \eta_{ij}$$

when student $i \in \mathcal{I}$ who is currently enrolled in middle school program $m(i)$ enrolls at high school program $j \in \mathcal{J}$. $\tilde{X}_j$ is the vector of observable characteristics of high school program j, Z_i^H is the vector of student observable characteristics when applying to high schools (for example, test scores which may depend on $m(i)$), and $\tilde{d}_{ij}$ is the distance between student i's residence and program j's location. γ_i^H is the vector of student i's unobserved taste on $\tilde{X}_j$, and η_{ij} is idiosyncratic preference shock that is iid for each i and j.

2.4.1.2 Period 1: Middle School Application

Forward-Looking Behavior Each student is forward-looking. Therefore, in the first period, she takes into account that enrolling in a given middle school program may affect her payoffs in the second period. Although we can abstract from modeling students' strategy profiles using ex-post stability, we need to model how students form expectations on the continuation value of choosing some middle school program.

In a school choice situation, each student is playing an incomplete information game. The uncertainties that affect a student's payoff are other students' types and the realization of lottery tie-breakings. To the extent that ex-post stability implies that each student is assigned to her most preferred school among the set of schools at which her score is above the ex-post cutoff in scores i.e., feasible, the ex-ante uncertainties in the cutoffs are sufficient statistics of the uncertainties present in the economy that affect students' payoffs. Define a random vector $\omega \in \Omega$ that captures the high school programs' cutoffs in terms of *ex-post* scores with some distribution $H(\omega)$. ω is unobserved ex-ante to every player in the game, and ex-post determines the assignments. Across ω, V_{ij} is invariant but the probability of a high school program being feasible to the student varies, and thus ω affects the expected payoff from high school choice. Denote the set of high school programs feasible to student i given the cutoff realization ω, attendance of middle school program m, and her own characteristics Z_i^H (which may depend on m) by $O_i(Z_i^H, m; \omega)$. Note that $O_i(Z_i^H, m; \omega)$ is explicitly a function of Z_i^H and middle school attendance m, capturing the aforementioned *priority* channel.[27]

In the first period, each student forms an expectation of the utility she will get from the second stage after she attends m, conditional on the state variables in the first period. Using ex-post stability, student i who attended m will be assigned to the high school program that gives her the maximum utility among those in $O_i(Z_i^H, m; \omega)$ given ω, and the payoff equals $\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij}$. Table 2.6 summarizes what is known to student i in each period.[28]

	Unobserved Taste on School Char.		Idiosyncratic Preference Shock		Program Characteristics	Student's own Characteristics	Ex-post Cutoff of High Schools
	γ_i^M	γ_i^H	ϵ_{im}	η_{ij}	$X_m, \tilde{X}_j$	Z_i^M, Z_i^H	ω
1st Period (Middle School Application)	✓		✓		✓	✓	
2nd Period (High School Application)	✓	✓	✓	✓	✓	✓	

Table 2.6: State Variables in Each Period

In the first period (middle school application), students submit ROLs on middle

[27]Recall that the priority channel includes two possible effects of a given middle school. First, the change of test scores which can influence a student's standings at programs that actively screen applicants based on test scores, and second, the change of eligibility or priority group. The former is captured by Z_i^H, and the latter is captured by the additional inclusion of m in the notation.

[28]We assume high school program characteristics are exogenous and fixed which are known to students in the first period. This is supported by the fact that school characteristics are stable over the years. Also, we assume a student has perfect foresight on what Z_i^H she will have by attending m . section 2.6.4 provides details on how we estimate each middle school's production function of Z_i^H using a value-added model.

school programs satisfying ex-post stability, based on the utilities:

$$U_{im} = \underbrace{u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im}}_{\text{Flow utility of attending } m} + \underbrace{\delta\, E_{\gamma_i^H, \omega, \eta_i, Z_i^H}\left[\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij} \middle| Z_i^M, \gamma_i^M, \epsilon_i, m\right]}_{\text{Continuation value of attending } m}$$

when student $i \in \mathcal{I}$ enrolls at middle school program $m \in \mathcal{M}$. X_m is the vector of observable characteristics of middle school program m, Z_i^M is the vector of student observable characteristics when they apply to middle schools, and d_{im} is the distance between student i's residence and program m's location. γ_i^M is the vector of student i's unobserved taste on X_m, and ϵ_{im} is an idiosyncratic preference shock that is iid for each i and m.

U_{im} consists of two parts: the flow utility of attending m and the continuation value of attending m. The continuation value of attending m is the expectation of the maximum utility of among those in $O_i(Z_i^H, m; \omega)$, where the expectation is with respect to the state variables in the second period (including ω) that are unknown to the student in the first period, and conditional on the state variables known in the first period as well as the middle school program m.

Unobservables and Expected Utilities We assume the following relationships on the unobservables.

$$\eta_{ij} \perp \epsilon_{im}, \forall i, j, m \tag{2.4}$$

$$\gamma_i^H \perp \eta_{ij}, \forall i, j \tag{2.5}$$

$$\gamma_i^M \perp \epsilon_{im}, \forall i, m \tag{2.6}$$

$$\omega \perp (\gamma_i^H, \eta_{ij}) \big| X_m, Z_i^M, \gamma_i^M, \epsilon_i, m, \forall i, j, m \tag{2.7}$$

The first three are standard. The first assumption states the idiosyncratic preferences in each period, ϵ_{im} and η_{ij}, are independent for all i, j, m. The second and third assumptions state that fixing a period, the unobservable tastes on program characteristics are independent of the idiosyncratic preferences.

Finally, the fourth assumption states that the uncertainty in the cutoff is independent of the unobservable tastes on high school program characteristics as well as the idiosyncratic preferences in the second period, conditional on the state variables in the first period and middle school attendance m. This assumption is valid as long as the economy is large enough so that each student acts like a 'price-taker' and cannot affect the cutoffs of high schools.

Given the assumptions, we can rewrite U_{im} as

$$U_{im} = u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) + \epsilon_{im} \tag{2.8}$$

$$+ \delta \int_\omega E_{\gamma_i^H, \eta_i}\left[\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij} \middle| Z_i^M, \gamma_i^M, m \right] dH(\omega | Z_i^M, \gamma_i^M, \epsilon_m, m) \tag{2.9}$$

2.4.2 Estimation

Parameterization: Preferences We parameterize the payoff functions using a random coefficient model.

First, the flow utilities in each period are:

$$\begin{aligned} u\left(X_m, Z_i^M, d_{im}, \gamma_i^M\right) &= \tilde{u}\left(X_m, Z_i^M, \gamma_i^M\right) - \lambda^M d_{im} \\ &= X_m' \beta_i^M - \lambda^M d_{im} \\ v\left(\tilde{X}_j, Z_i^H, \tilde{d}_{ij}, \gamma_i^H; m(i)\right) &= \tilde{v}\left(\tilde{X}_j, Z_i^H, \gamma_i^H; m(i)\right) - \lambda^H \tilde{d}_{ij} \\ &= \tilde{X}_j' \beta_i^H - \lambda^H \tilde{d}_{ij} \end{aligned}$$

where λ^M and λ^H capture the disutility of traveling, and we allow students' tastes on program observable characteristics to be heterogeneous across i, as captured by β_i^M, β_i^H. We normalize the location of the utilities by setting $\tilde{u}(\cdot) = \tilde{v}(\cdot) = 0$ if all of their arguments are equal to zero. Additionally, we assume

$$\begin{aligned} (\gamma_i^M, \epsilon_{im}) \perp d_{im} &\Big| X_m, Z_i^M \\ (\gamma_i^H, \eta_{ij}) \perp \tilde{d}_{ij} &\Big| \tilde{X}_j, Z_i^H, m(i) \end{aligned}$$

which together with the additive separability of distances $d_{im}, \tilde{d}_{ij}$ provide nonparametric identification of the utilities $\tilde{u}$ and $\tilde{v}$ (Agarwal and Somaini, 2018b).

Let the dimension of X_m, $\tilde{X}_j$ and consequently that of β_i^M, β_i^H be L. For the l-th program characteristic, we assume that the random coefficients are parametrized as:

$$\beta_{i,l}^M = \beta_{0,l}^M + Z_i^{M\prime}\beta_{Z,l}^M + \gamma_{i,l}^M$$
$$\beta_{i,l}^H = \beta_{0,l}^H + Z_i^{H\prime}(m(i))\beta_{Z,l}^H + \underbrace{\sum_{\tau=1}^{T}\rho_{\tau,l}1\left(\tau(m(i)) = \tau\right)}_{\text{Middle school type effect}} + \gamma_{i,l}^H$$

$\beta_{0,l}^M$, $\beta_{0,l}^H$ capture the common valuation of all students on the l-th program characteristic in each period. The interaction terms $Z_i^{M\prime}\beta_{Z,l}^M$ and $Z_i^{H\prime}\beta_{Z,l}^H$ allow individual tastes to depend on individual observable characteristics Z_i^M and Z_i^H respectively. Note that the student observable characteristics $Z_i^H(m(i))$ when applying to high schools are allowed to be dependent on the student's middle school $m(i)$.

Importantly, $\sum_{\tau=1}^{T}\rho_{\tau,l}1(\tau(m(i)) = \tau)$ is what we call the *middle school type effect*, where $\tau(m(i))$ is the type of i's attended middle school $m(i)$. It allows students who attends middle schools with some type $\tau = 1, \cdots, T$ to have a different mean valuation of program characteristics. $\rho_{\tau,l}$ plays a similar role as the treatment effect β in Equation 2.1 when the outcome variables are the characteristics of the programs students applied to.

$\gamma_i^M = (\gamma_{i,1}^M, \cdots, \gamma_{i,L}^M)$ and $\gamma_i^H = (\gamma_{i,1}^H, \cdots, \gamma_{i,L}^H)$ capture students' unobservable tastes on middle school and high school program characteristics. They are serially correlated, which generates a source of sorting across two periods. We assume:

$$\gamma_i^H = \underbrace{diag(\rho_0)\gamma_i^M}_{\text{serial correlation}} + \xi_i \tag{2.10}$$

or

$$\begin{pmatrix} \gamma^H_{i,1} \\ \vdots \\ \gamma^H_{i,L} \end{pmatrix} = \begin{pmatrix} \rho_{0,1} & & 0 \\ & \ddots & \\ 0 & & \rho_{0,L} \end{pmatrix} \begin{pmatrix} \gamma^M_{i,1} \\ \vdots \\ \gamma^M_{i,L} \end{pmatrix} + \begin{pmatrix} \xi_{i,1} \\ \vdots \\ \xi_{i,L} \end{pmatrix}$$

ξ_i captures the innovation on the unobservable tastes that is only realized in the second period. We assume that $\gamma_i^M \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\gamma)$, $\xi_i \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\xi)$ and they are mutually independent, where we allow Σ_γ and Σ_ξ to be fully flexible. We impose a restriction that $diag(\rho_0)$ is a diagonal matrix implying that the unobservable taste on one middle school program characteristic ($\gamma^M_{i,l}$) does not impact the unobservable taste on other high school program characteristics ($\gamma^H_{i,l'}, \forall l' \neq l$).[29]

Finally, we assume that the idiosyncratic preferences ϵ_{im} and η_{ij} both follow Extreme Value Type-I (EVT1) distribution. Together with the assumption in (2.5), it implies that a part in the continuation value expression can be further simplified to

$$\begin{aligned} &E_{\gamma_i^H, \eta_i}\left[\max_{j \in O_i(Z_i^H, m; \omega)} V_{ij} \middle| Z_i^M, \gamma_i^M, m \right] \\ &= \int_{\eta_i} \left(\mu + \log \left(\sum_{j \in O_i(Z_i^H, m; \omega)} \exp(V_{ij}) \right) \right) d\eta_i \end{aligned}$$

which helps computation dramatically.

Source of Identification We provide informal identification arguments.

Our primary identification concern is to distinguish the causal effect of the type of middle school on tastes on high schools ($\{\rho_\tau\}_\tau$) from students' unobservable tastes on those (γ_i^H). In the data, we observe a large correlation between the high school characteristics a student applies and is assigned to and the middle school characteristics she attends (see Table 2.1). A large part of this relationship can be

[29]Even with this restriction, the arbitrary correlation among γ_i^M we allow enables γ_i^M to have implicit impact on γ_i^H throught the correlation structure.

explained by students' observable characteristics that are constant, but even conditional on observable characteristics, there is still positive correlation (see Table 2.5). This could be attributable to either the consistency of the individual student's unobserved tastes over time or the treatment effect of attending a particular type of middle school. The key to distinguishing between these sources of explanation comes from the panel structure of the data. We observe each student's middle school and high school ROLs. ρ_0 is identified by the degree to which the same student's middle school and high school applications look similar, after controlling for her observable characteristics.

Meanwhile, ρ_τ is identified by how similar the high school applications are across students attending middle schools of the same type. Notably, we implicitly rely on the quasi-random variation in school assignments generated by the tie-breaking rule. The quasi-random assignment generates variation in what type of middle school a student attends beyond her middle school application. Without the quasi-random tie-breaking, observably similar students' attending different middle schools would be all attributable to the difference in γ_i^M once we assume nonparametric identification of the unobserved taste γ_i^M. Thanks to the quasi-random variation, we have variations in which type of school a student attends beyond what can be explained by students' observable characteristics and unobserved tastes. The remaining variation in the application that is explained by neither within-student consistency nor across-student (of the same middle school) correlation is captured in Σ_ξ.

Estimation We focus on students and residents of Staten Island for tractability. Staten Island can be effectively treated as an independent school district in NYC since commuting outside of Staten Island is very costly for students.[30] Indeed, a majority of students who reside in Staten Island and enrolled in middle schools in Staten Island apply to high school programs only in Staten Island. For example,

[30]One can travel from Staten Island to other boroughs in NYC only via the Staten Island Ferry or the Verrazzano-Narrow Bridge, the only ground transportation route to Brooklyn. See Figure 2.7 for the map of NYC school districts.

in 2017-18 high school applications, only 1.8% ever ranked a high school program outside of Staten Island among Staten Island middle school students. There are 2,626 students applying to 20 middle school programs (14 schools) and 47 high school programs (10 schools) in our estimation sample.[31]

For the program characteristics X_m and $\tilde{X}_j$, we use three variables: fraction of high performers (current 6th (9th) graders who belong to the top 1/3 in terms of the average of statewide ELA and math exams), fraction of White students in current 6th (9th) grade, and if the program focuses on STEM-related fields. For the student charateristics, we use Asian, Black, Hispanic, Free of Reduced Lunch (FRL) status, English Language Learner (ELL) status, and the mean of most recent math and ELA standardized test scores (normalized to mean 0 and std 1). Finally, we include two types of middle schools, high achievement middle schools and high minority middle schools, defined in the same way as in section 2.3.

We assume there is a common value of outside options 0_m and 0_h for middle school programs and high school programs, respectively. That is,

$$U_{i0_m} = \vartheta_{mi} + \epsilon_{i0_m}$$
$$V_{i0_h} = \vartheta_{hi} + \eta_{i0_h}$$

where $\epsilon_{i0_m}, \eta_{i0_h}$ both follow EVT1. subsection 2.6.4 provides more details on the procedure we use to estimate our model.

To interpret the data, we employ ex-post stability as the identifying assumption. Ex-post stability essentially enables us to interpret the school choice data as a conditional multinomial logit model, where a student's *choice* is the assigned program, and the *choice set* is the set of programs that were feasible to the student.[32] Traditionally, weak truth-telling (WTT) (or strict truth-telling (STT) in

[31]Staten Island is on average a richer borough with more White and slightly higher performing students compared to the rest of the New York City. The fraction of subsidized lunch status was about 54% (72% citywide), the fraction of White students was about 56% (17% citywide), and the average statewide Math exam score was 315 (311 citywide) in academic year 2014-2015.

[32]Note that in any feasible set, the outside options 0_m and 0_h are included respectively for middle school and high school choice. Also, the exogeneity of choice set is satisfied with the large market assumption i.e., when each student cannot affect the cutoffs.

the presence of outside options) has been widely adopted in the empirical school choice literature as an assumption on student's behavior to interpret the ROL data (for example, Abdulkadiroğlu et al., 2017a) based on strategyproofness of SPDA. However, the truth-telling assumption can be problematic when deviations from truth-telling (often regarded as *mistakes* in the literature) do not affect a student's payoff. For example, a low performing middle school senior may not choose to apply to a highly competitive screened high school program because she knows that there is zero chance of admission. This in turn does not affect her payoff, even if it is one of her most desirable programs. Therefore, we use ex-post stability as our preferred assumption, which is robust to payoff-irrelevant mistakes as in the example. We also estimate the model using STT as an additional robustness check in subsection 2.6.5.

2.4.3 Results

2.4.3.1 Model Estimates

We estimate via maximum simulated likelihood estimation (MSLE) using sparse grids quadrature (Heiss and Winschel, 2008).[33] Table 2.7 provides the model estimates on our main specification. The model estimates have mainly four implications.

First and most importantly, we reconfirm that middle schools affect how students value different high school characteristics. All else equal, attending a high achievement middle school makes a student will to travel 0.11 miles (resp., 0.31 miles) more for a 10 pp increase in the fraction of high performers (resp., the fraction of White students). On the other hand, attending a high minority middle school makes a student will to travel 0.17 miles more (resp., 0.28 miles less) for a 10 pp increase in the fraction of high performer (resp., the fraction of White

[33]See section 2.6.4 for the description of the likelihood function.

students).[34] [35] [36] [37]

Second, students value the continuation value of attending some middle school program, as shown with the positive estimate on δ. The estimate suggests that students' willingness to travel for 1 standard deviation increase in the continuation value across middle school programs equals 0.81 miles. That is, they value middle school programs that enable one to enjoy a higher utility in high school applications.

Third, serial correlation of unobservable tastes on the program characteristics exists, as implied by the estimate of ρ_0, in particular by the positive relationship between the unobservable taste on the fraction of White students in middle school and high school applications. The estimates of ρ_0, Σ_γ, and Σ_ξ imply that the proportion of variance of γ_i^H, the unobservable taste on high school characteristics, which can be attributed to the variance of γ_i^M, the unobservable taste on middle school characteristics, is 18.45% and 28.59% for the fraction of high performer, and the fraction of White students. It suggests that students' selection based on unobservable

[34]The conversion into willingness to travel is done by dividing the coefficient of interest by the coefficient on distance. For example, attending a high achievement middle school makes a student will to travel $0.546/0.509/10 = 0.11$ miles more for 10 pp increase in the fraction of high performer. The average commuting distance to each assigned high school in the data is 2.3 miles.

[35]The estimates of the effect of attending high minority middle school provide a potential explanation for the nearly null effect we find in Table 2.19. In reality, the fraction of high performing students and the fraction of White students are positively correlated ($r = 0.62$ among Staten Island high school programs), making the effects of high minority middle schools on the taste on high schools cancel out each other. This results in nearly null treatment effects of high minority middle schools since we do not consider each program's characteristics simultaneously in the two-stage least squares in section 2.3.

[36]We find these middle school effects are robust to allowing heterogeneous effects by race group. For White/Asian students, attending a high achievement middle school makes them will to travel 0.12 miles (resp., 0.41 miles) more for a 10 pp increase in the fraction of high performers (resp., the fraction of White students) and attending a high minority middle school makes them will to travel 0.24 miles more (resp., 0.34 miles less) for a 10 pp increase in the fraction of high performer (resp., the fraction of White students). For Black/Hispanic students, attending a high achievement middle school makes them will to travel 0.12 miles (resp., 0.08 miles) more for a 10 pp increase in the fraction of high performers (resp., the fraction of White students) and attending a high minority middle school makes them will to travel 0.11 miles more (resp., 0.22 miles less) for a 10 pp increase in the fraction of high performer (resp., the fraction of White students).

[37]These estimates imply that middle school segregation has a reinforcing effect on high school segregation. In terms of academic segregation, students who attended high quality middle schools aspire to higher quality high schools. In terms of racial segregation, attending middle schools with many students of the same race strengthens racial homophily.

tastes across two periods is present.[38] [39]

Discussion on other preference estimates follows.

First, overall, students prefer programs with a higher fraction of high performers and White students. Whether a middle school program is STEM does not significantly change students' preferences, and students slightly dislike STEM high school programs.

Second, students have heterogeneous preferences on program characteristics based on observable student characteristics, while the degree of heterogeneity is much smaller for middle school programs. For example, the preference for the fraction of high performers is stronger if a student herself has higher baseline test scores, and the preference for the fraction of Whites is much smaller for Black or Hispanic students than White students.

Third, students dislike commuting, which is captured by positive estimates on λ_m and λ_h.

2.4.3.2 Goodness of Fit

We evaluate how well the model fits the observed data by comparing measures calculated using the data to those calculated using the simulations based on model estimates. In Table 2.8, we calculate the average characteristics of *assigned schools* for each type of *students* and the average characteristics of *assigned students* for each type of *schools*.

[38]That is, we calculate $\frac{Var(\rho_{0,l}\gamma^M_{i,l})}{Var(\gamma^H_{i,l})} = \frac{Var(\rho_{0,l}\gamma^M_{i,l})}{Var(\rho_{0,l}\gamma^M_{i,l})+Var(\xi_{i,l})}$ for each $l = 1, \cdots, L$. It has a similar interpretation as the R^2 measure in standard least squares regressions.

[39]The random taste (the student's 'residual' heterogeneous taste that is left after controlling for student's characteristics) is weak for some school characteristics whose variances are estimated with small significance. We interpret this as showing that the student characteristics we included are rich enough to capture low heterogeneity.

	Middle Schools			High Schools		
	est	se		est	se	
Panel A: Preference Estimates						
Fraction of High Performer						
Main Effect	4.944	(1.144)	***	0.795	(0.272)	***
Asian	-1.267	(1.947)		0.827	(0.390)	**
Black	6.820	(1.961)	***	-0.199	(0.462)	
Hispanic	1.781	(1.288)		-0.275	(0.330)	
Free or Reduced Lunch	-0.881	(1.130)		-0.922	(0.271)	***
English Language Learner	-1.804	(2.309)		0.342	(1.177)	
5th Grade Test Score	1.088	(0.581)	*	1.652	(0.141)	***
Fraction of White						
Main Effect	3.056	(0.875)	***	4.931	(0.343)	***
Asian	0.976	(1.588)		-2.011	(0.599)	***
Black	-6.444	(1.721)	***	-1.520	(0.613)	**
Hisp	-1.666	(1.047)		-1.060	(0.421)	**
Free or Reduced Lunch	-0.565	(0.886)		0.162	(0.346)	
English Language Learner	0.752	(1.954)		-0.24	(1.202)	
5th Grade Test Score	-0.951	(0.468)	**	0.341	(0.126)	***
1(STEM)						
Main Effect	0.281	(0.198)		-0.676	(0.123)	***
Asian	0.157	(0.324)		-0.174	(0.200)	
Black	-0.420	(0.269)		0.090	(0.196)	
Hisp	0.121	(0.213)		0.083	(0.144)	
Free or Reduced Lunch	-0.122	(0.198)		0.257	(0.126)	**
English Language Learner	0.062	(0.345)		1.005	(0.326)	***
5th Grade Test Score	-0.159	(0.096)	*	0.003	(0.044)	
Panel B: Middle School Type Effects						
Type 1 (High Achievement MS)						
Fraction of High Performer				0.546	(0.276)	**
Fraction of White				1.600	(0.318)	***
1(STEM)				-0.322	(0.137)	**
Type 2 (High Minority MS)						
Fraction of High Performer				0.875	(0.301)	***
Fraction of White				-1.447	(0.378)	***
1(STEM)				0.198	(0.136)	
Panel C: Unobservable Tastes						
ρ_0				0.074	(0.044)	*
				0.429	(0.127)	***
				-0.035	(0.118)	
(1,1) of Σ_γ	18.461	(10.853)	*			
(1,2)	-17.930	(9.653)	*			
(1,3)	-0.186	(1.626)				
(2,2)	23.168	(10.222)	**			
(2,3)	2.765	(2.018)				
(3,3)	1.163	(0.697)	*			
(1,1) of Σ_ξ				0.447	(0.316)	
(1,2)				-2.184	(0.950)	**
(1,3)				0.411	(0.163)	**
(2,2)				10.670	(2.877)	***
(2,3)				-2.006	(0.512)	***
(3,3)				0.377	(0.193)	*
Panel D: Other Parameters						
Outside option	2.698	(0.367)	***	-0.371	(0.175)	**
Distance	0.655	(0.038)	***	0.509	(0.018)	***
Discount Factor	0.877	(0.064)	***			

Notes: We report the preference estimates of the main model described in section 2.4. School characteristics 'Fraction of High Performer' and 'Fraction of White' are between 0 and 1, and '1(STEM)' is an indicator variable. In Panel A, Main Effect is the common taste (β_0^M, β_0^H), and we also include interactions of each school characteristics with Asian, Black, Hispanic, Free or Reduced Lunch (FRL) status, English Language Learner (ELL) status, 5th Grade Test Score in the following rows (β_Z^M, β_Z^H). Robust standard errors are reported in parentheses.

Panel A. Average Characteristics of Assigned Schools by Student Characteristics								
	Middle Schools				High Schools			
	% High Performers		% Black/Hisp		% High Performers		% Black/Hisp	
	Data	Model	Data	Model	Data	Model	Data	Model
Asian	36%	37%	40%	39%	33%	33%	43%	43%
Black	27%	32%	66%	59%	25%	24%	62%	63%
Hispanic	31%	34%	53%	49%	29%	27%	52%	55%
White	45%	45%	24%	25%	39%	37%	30%	34%
English Language Learner (ELL)	27%	30%	59%	53%	24%	24%	63%	62%
Free/Reduced Lunch (FRL)	35%	37%	45%	43%	31%	30%	48%	49%

Panel B. Average Characteristics of Assigned Students by School Type								
	Middle Schools				High Schools			
	High Achievement		High Minority		High Achievement		High Minority	
	Data	Model	Data	Model	Data	Model	Data	Model
Asian (%)	9%	9%	9%	9%	9%	9%	7%	8%
Black (%)	4%	3%	25%	24%	4%	4%	30%	23%
Hispanic (%)	12%	12%	41%	39%	18%	15%	42%	40%
White (%)	74%	75%	25%	27%	68%	71%	20%	28%
English Language Learner (ELL) (%)	2%	1%	10%	9%	3%	3%	11%	9%
Free/Reduced Lunch (FRL) (%)	41%	40%	77%	73%	46%	44%	78%	74%
5th Grade Math Score	322.6	322.6	304.2	307.5	320.0	322.3	301.9	303.7
From High Achievement MS (%)					57%	61%	10%	9%
From High Minority MS (%)					10%	10%	62%	49%

Notes: For model based simulations, we report the average result from 5,000 DA simulations based on model estimates (100 draws of unobservables $\times$ 50 draws of lotteries). The definitions of 'high achievement' and 'high minority' are as described in subsection 2.2.2. The scale of 5th grade math score is from 125 to 402.

Table 2.8: Goodness of Fit

We find the measures based on model simulations well match those based on the observed data and hence, our dynamic model can be credibly used to predict the impacts of counterfactual policies in subsection 2.4.5.

First, in Panel A, we find that the average characteristics of the assigned schools for each type of student are very similar across data and model simulations, both for middle schools and high schools. For example, in the data, Asian students on average are assigned to middle (resp., high) schools with the fraction of high performers equal to 36% (resp., 33%) and the fraction of Black of Hispanic students equal to 40% (resp., 43%). Using the model estimates, we predict such students are on average assigned to middle (resp., high) schools with 37% (resp., 33%) in terms of the fraction of high performers and 39% (resp., 43%) in terms of the fraction of Black or Hispanic students. Second, in Panel B, the distributions of student observable characteristics at each type of schools are also very similar across data and model simulations. For example, our model almost perfectly predicts the racial composition at each type of middle and high schools. Importantly, in the last two rows of Panel B, our model predicts the transition from each type of middle schools to each type of high schools reasonably well.

Additional goodness of fit measures such as how well the model predicts the assignments, and how well the model predicts students' revealed preferences are available in subsection 2.6.6.

2.4.4 Decomposition of Effects of Middle Schools

Recall that the model allows two channels of middle school effects on high school assignments: the *application* channel and the *priority* channel. To illustrate the relative importance of the two, we perform the following illustrative exercise: what happens if we exogenously make a student who is currently attending a 'bad' middle school attend a 'good' middle school?

To this end, we randomly select students (10% of the entire sample), and counterfactually assign them to middle 'school A' with the lowest average test score in Staten Island as a benchmark (**Counterfactual 0**). Next, for each student, we counterfactually change their middle school enrollment to another middle 'school B' with the highest average test score in Staten Island (**Counterfactual 1**) *one student at a time*, and make them apply to high schools in the following alternative scenarios.[40]

1. **Full**: both *application* and *priority* channels are active.
2. **Application**: shut down the *priority* channel. That is, we do not allow students' priorities to change at the middle school level.
3. **Priority**: shut down the *application* channel. That is, we do not allow students' tastes on high school programs to change at the middle school level.

[40] Instead of alternatively assigning students to a hypothetical middle school, we choose among existing middle schools in Staten Island to ensure we have a realistic estimate of effects of rearranging students' middle school assignments.

We keep track of how the students' high school assignments change compared to when they attend middle school A (**Counterfactual 0**) in each scenario. We seek to evaluate of the effect in **Full** (the total effect of exogenously changing middle schools), how much can be explained by the *application* channel (the effect in **Application**), or by the *priority* channel (the effect in **Priority**).[41] This procedure treats each student essentially as a 'price-taker' who takes the current equilibrium as given and considers what will be the change in her high school assignment when *only* her middle school enrollment is exogenously changed. Also, randomly selecting students and exogenously assigning them to a benchmark school enable us to be free of sorting of students into middle schools based on unobservables. Note that in these regards, the measures we report have an interpretation as the average treatment effect (ATE) of changing middle schools.

	Middle School Types		Average Test Score
	High Achievement?	High Minority?	
School A (Counterfactual 0)		✓	602.01
School B (Counterfactual 1)	✓	✓	611.42

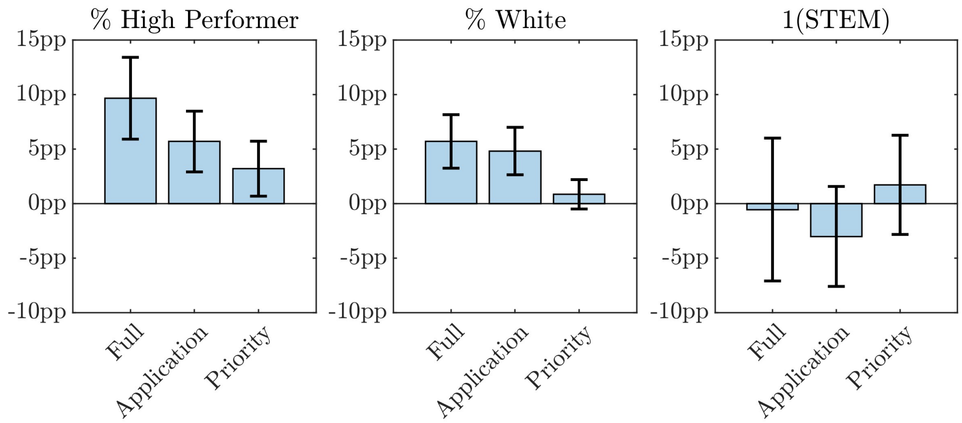
Notes: For Counterfactual 0, we choose the middle school which has the lowest average test score, whose category is 'low achievement and high minority'. For Counterfactual 1, we choose the middle school that has the highest average test score, whose category is 'high achievement & low minority'. Results for alternative type of counterfactual 'good' middle schools are reported in section 2.6.8. Average test scores are the average of 8th grade statewide test scores of current seniors. The scale is from 500 to 650.

Table 2.9: Alternative Assignment to Middle Schools

Figure 2.2 reports the results.[42] We find that the impact of attending a different middle school through the *application* channel dominates the *priority* channel. For example, when a student attends a 'good' middle school B instead of a 'bad' middle

[41]Based on the preference estimates in Table 2.7, we simulate 100 draws of unobservables and 50 draws of tie-breaking lotteries, and run SPDA for each randomly selected student in each scenario. Hence for each student, we have 100 counterfactual application lists and $100 \times 50 = 5,000$ counterfactual assignments for each scenario. We report the measures by taking the average over the unobservable draws, lottery draws, and the randomly selected students. The mean standard deviation across students is used to construct 95% confidence intervals.

[42]See section 2.6.8 for additional figures and tables.



Notes: We report the decomposition of middle school effects on high school assignments using the model estimates in Table 2.7. We use a randomly selected subsample of student (10% of the entire sample), and counterfactually assign them to Counterfactual 0 middle school. Then we calculate the average change in the characteristics of the assigned high school program when they are counterfactually assigned to Counterfactual 1. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and SPDA are run, resulting in 5,000 simulated assignments. The bar graphs and 95% confidence intervals are plotted using the average (across unobservables and lottery draws) of mean and standard deviation across students. The corresponding numbers are calculated in Table 2.21

Figure 2.2: Decomposition of Effects of Middle Schools on High School Assignments

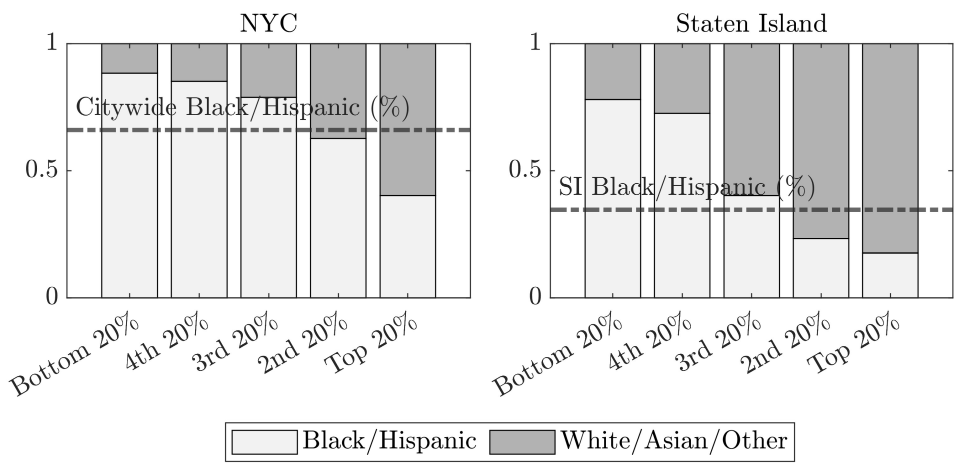
school A, we find that about 59% of the total effect on the assigned high school's fraction of high performers can be explained by the *application* channel, while the *priority* channel can only explain about 33%. When it comes to the fraction of White students, the *application* channel explains about 84% and the *priority* channel explains about 15% of the total effect.

The decomposition exercise reconfirms that middle schools play an important role in high school choice outcomes and that the effect mainly occurs by affecting how students apply to high schools. The result motivates into the possibility of changing high school assignments through the change of students' applications by changing *middle school assignments*, which is explored in the next section.

2.4.5 Counterfactual Analysis

2.4.5.1 Segregation in NYC Public High Schools

NYC high schools are intensely segregated. Figure 2.3 plots the racial composition of high schools by quintiles of average performance of enrolled students. It shows that Black and Hispanic students are underrepresented at high perfoming high schools, while they are overrepresented in low performing high school programs.



Notes: We plot the racial composition of high school programs by the quintiles of average performance level of students. For a given program, the fraction of 9th graders in AY 2017-18 whose average of 8th grade statewide Math and ELA scores fall in the first tercile are calculated, and then the high school programs are classified into each quintile of it. The left panel uses the entire NYC high schools, and the right panel uses the Staten Island high schools. The fraction of Black and Hispanic students in each sample is plotted in the grey dotted line.

Figure 2.3: High School Racial Composition by Performance Level

The NYC government has long acknowledged this problem. Most recently, partially due to the cancellation of statewide test scores due to the spread of SARS-CoV-2 during 2020 and onwards, the city announced changes to the NYC public school system to deal with racial segregation.

> Mayor Bill de Blasio announced on Friday major changes to the way hundreds of New York City's selective middle and high schools admit their students. [...] Black

> and Latino students are significantly underrepresented in selective middle and high schools. [...] The city will eliminate all admissions screening for the schools for at least one year [...] New York will also eliminate a policy that allowed some high schools to give students who live nearby first dibs at spots. — The New York Times[43]

Motivated by NYC's intervention, we perform an illustrative counterfactual analysis. Namely, we eliminate all priority rules currently employed by the schools, including selecting students based on test scores and any form of geography-based priority rules including zoned programs.[44] Any remaining racial imbalance would be due to students' applications. We evaluate the following three alternative interventions:

1. **MS**: we only get rid of priority rules of middle schools.
2. **HS**: we only get rid of priority rules of high schools.
3. **MSHS:** we get rid of priority rules of both middle and high schools.

In each scenario, we solve the new equilibrium based on the model estimates and compare how students' high school assignments change compared to the status quo (**Current**).

To this end, we use the preference estimates obtained in Table 2.7, and simulate 100 independent draws of the unobservables and 50 independent draws of tie-breaking lotteries, and run the SPDA algorithm for each pair of unobservable draw and lottery draw, obtaining 5,000 counterfactual assignment results for each counterfactual scenario. In all results we present, we report the average of each

[43] Shapiro, Eliza, N.Y.C. to Change Many Selective Schools to Address Segregation, The New York Times, 18 December 2020.

[44] Ideally, we would evaluate the exact policy employed by NYC. However in our Staten Island sample, there is only one Screened middle school program, and the geographic priority rules of high school programs are entirely based on if a student is a resident or student of Staten Island, making virtually all students in our sample unchanged by employing the actual policy change. Instead, we remove priority rules altogether so that schools admit students solely based on lotteries, and hence assignments are entirely decided by how students apply to schools. This choice of policy intervention highlights the role of how students submit their choices, which is closely related to the main findings of this paper that middle schools mainly affect how students submit choices in subsequent school choice.

measure across the pairs of unobservable draw and lottery draw for the three counterfactual scenarios—**MS**, **HS** and **MSHS**. To calculate the measures under the current regime (**Current**), we use the observed assignment results in the data.

Given the importance of middle schools in how students apply and are assigned to high schools, we have two conjectures.[45] First, even under **MS**, which high school programs students are assigned to will change due to the change in middle school assignments and hence the applications to high school programs through the *application* and *priority* channel. Second, the effects of intervention on high school assignments under **MSHS** will be stronger than under **HS**, since **HS** only reforms the 'supply' side while **MSHS** reforms both the 'demand' and 'supply' sides of school choice.

2.4.5.2 Predictions

We evaluate the impacts of counterfactual policy changes along two dimensions: the impacts on minority students' school assignments and on overall segregation in schools. We assume that the school characteristics are fixed as under the status quo (**Current**), which gives us the interpretation of the predictions as the short-term impacts.

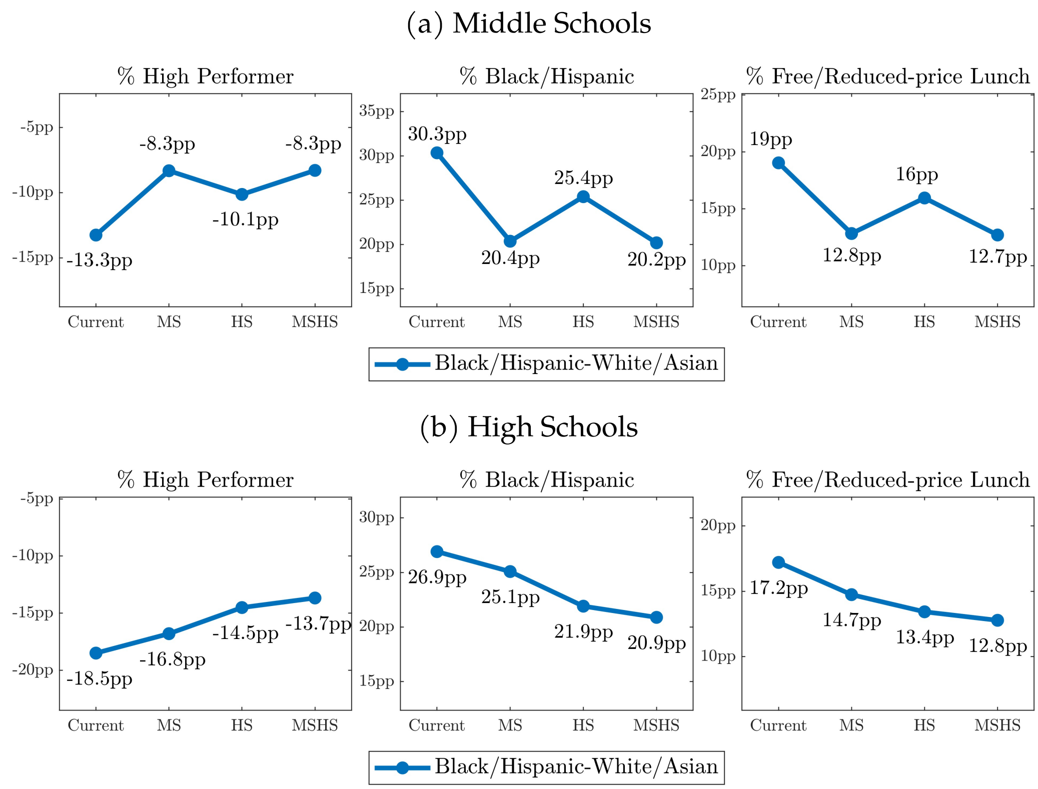
Effects on Minority Students' Assignments Figure 2.4 plots the relative difference of characteristics of assigned schools of Black or Hispanic students to those of White or Asian students.[46]

We largely confirm the conjectures. First, we find that intervening only at the middle school level (**MS**) alone can reduce the racial gap in high school assignments, amounting to about 40% of what intervening only at the high school level (**HS**) can achieve. Second, the effects of combining both interventions at the middle school level and the high school level (**MSHS**) are stronger for high schools assignments

[45] One would also expect that due to the forward-looking behavior of students in our model, how students are matched with middle school programs may change even under **HS**.

[46] See section 2.6.8 for the effects on the peer characteristics of co-assigned students.

than only intervening at the high school level (**HS**).[47] These effects are driven by **MS** effectively desegregating middle schools as shown in Panel (a).



Notes: The graph plots the gap of the characteristics of assigned school programs between Black or Hispanic versus White or Asian students in each counterfactual scenario. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and SPDA are run, resulting in 5,000 simulated assignments for each counterfactual scenario where the draws are fixed across scenarios. The mean across unobservable and lottery draws are reported for **MS**, **HS** and **MSHS**.

Figure 2.4: Racial Gap in Characteristics of Assigned School Programs in Staten Island

[47] However the marginal gain of **MS**→**MSHS** (**HS**→**MSHS**) is smaller than that of **Current**→**HS** (**Current**→**MS**), suggesting a possible substitutability of those two policy interventions.

Effects on Aggregate Segregation Measures We calculate two measures in order to summarize the aggregate segregation.[48] First, we calculate a measure of racial segregation, known as the Theil's H index (Panel A of Table 2.10). Theil's H index calculates a measure of the evenness of ethnic groups across programs based on multigroup entropy scores. It varies between 0 and 1, where 0 means maximum integration and 1 means maximum segregation.

Second, we calculate the sorting indices for three student characteristics: 1(Black or Hispanic), baseline standardized test scores, and the median income of the census tract that a student resides in (Panel B of Table 2.10). Sorting by 1(Black or Hispanic) provides a measure of segregation by race, sorting by test scores of students provides a proxy of sorting by student ability, and sorting by median census tract income provides a proxy of sorting by income. Each sorting index is between 0 and 1, and is defined as the ratio of the between-program variance of each student characteristic to its total variance (Yang and Jargowsky, 2006; He et al., 2021). That is, it measures the fraction of variance of a variable that between-program differences can explain. Hence, 0 means maximum integration and 1 means maximum segregation.

Table 2.10 reports the effects of policy intervention on aggregate segregation measures. We find similar patterns as for the effects on minority students' assignments.

Policy Implication Our counterfactual analysis has the following policy implication. While most existing policies for desegregation focus on reforming the *supply side*, i.e., how schools select students, it is crucial to consider how we can influence the *demand side* i.e., how students apply to schools. Reforming the *demand side* might have received less attention because it is not clear how to intervene effectively. We found in subsection 2.3.2, subsection 2.4.3 that students' high school assignments are largely affected by which middle schools they attend, mainly by changing their applications to high schools. Also the counterfactual analysis showed that intervening in middle schools alone can achieve desegregation in not only middle schools but also high schools by changing how students apply to high schools. In addition,

[48]See subsection 2.6.7 for more details on the description of measures.

		(1) Current	(2) MS	(3) HS	(4) MSHS
Panel A: Racial Segregation Measure					
Theil's H Index	Middle schools	0.216	0.106	0.189	0.102
	High schools	0.207	0.191	0.148	0.135
Panel B: Sorting Indices					
Sorting by Race	Middle schools	0.299	0.173	0.266	0.168
	High schools	0.301	0.263	0.212	0.201
Sorting by Ability	Middle schools	0.162	0.040	0.075	0.037
	High schools	0.357	0.309	0.119	0.117
Sorting by Income	Middle schools	0.456	0.237	0.432	0.230
	High schools	0.346	0.300	0.262	0.242

Notes: The table calculates the aggregate segregation measures of schools in each counterfactual scenario. Panel A calculates the Theil's H index, and Panel B calculates the sorting indices by race, ability and income. 100 sets of unobservable variables $(\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij})$ are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and SPDA are run, resulting in 5,000 simulated assignments for each counterfactual scenario where the draws are fixed across scenarios. The mean across unobservable and lottery draws are reported for **MS**, **HS** and **MSHS**.

Table 2.10: Aggregate Segregation Measures in Staten Island

conditional on intervening at the high school level, there is still room for desegregation which could be achieved by additionally intervening at the middle school level. Hence our findings suggest that the policy intervention in large school districts should take place early enough and that reforming middle school admissions may be one way to do so by changing students' applications to high schools.

2.5 Conclusion

In this paper, we provided a novel, evidence-based dynamic framework of school choices. We showed that a student's middle school and high school choices are closely related to each other by using NYC public school choice data. First, we leveraged the quasi-random middle school assignments generated by the tie-breaking feature in DA to provide empirical evidence of middle schools' causal effects on high school applications and assignments. Next, based on the empirical findings, we developed and estimated a novel dynamic framework of middle school and high school choices. We showed that the effects of middle schools on high school choice mainly operate by changing students' applications to high schools. Finally, using the dynamic framework, we provided a new perspective on how to understand and address segregation across public schools. Segregation patterns in middle and high schools are closely related, and hence the policy intervention for desegregation in high schools should begin early enough and reforming middle school admissions may be one such tool.

Our findings suggest two avenues of future research. First, having confirmed the dynamic relationship between middle school choice and high school choice, we may further move on to directly test for the dynamic complementarity of those two human capital investments (Cunha and Heckman, 2007; Heckman, 2007). While a credible quasi-randomization at two points for a given individual is hard to find, the fact that students are exposed to centralized school choice multiple times opens an avenue for a suitable research design to test for dynamic complementarity. Second, one may explore ways to design a student assignment mechanism that considers the dynamic relationship of school choices to achieve more equitable outcomes. We leave these for future research.

2.6 Appendix

2.6.1 Data and Sample Restriction

Data Sources The main data used is the administrative data from the New York City Department of Education, focusing on the 8th grade cohort in the academic year 2014-2015. This cohort applied to middle schools in the academic year 2014-15, and to high schools in the academic year 2017-18. There are four sets of data used to construct information on the applicants. First, high school application (HSAP) data includes information on each round of the application process (ROL, rank, priority, eligibility etc.) related to high school application and standardized test scores information. Second, middle school application (MSAP) data includes similar variables as HSAP but for middle school applications. Third, yearly June biographic data includes more comprehensive biographic data of the students including ethnicity, gender, disability status as well as information on attendance and punctuality. Lastly, Zoned DBN data includes information about student's residence (census tract level)[49] and which elementary, middle and high schools the students are zoned to. There exists a unique student ID variable that enables merging all dataset.

School information is constructed using the 2014-15 NYC Middle School Directory and 2017-18 NYC High School Directory that are published every year before the application process starts. This includes each program's previous year's capacity and the number of students who applied in the previous year, admission criteria (eligibility and priority), accountability data such as progress report, graduation rate and college enrollment rate, types of language classes provided etc. Other variables about current 6th graders in middle schools and 9th graders in high schools such as the composition of ethnicity or the fraction of high performing students are constructed using the previous year's student-level data.

NYC School Admission Methods Middle school programs use a variety of admission methods—Unscreened, Limited Unscreened, Screened, Screened: Language, Zoned and Talent Test. Unscreened programs admit students by a random lottery number, and Limited Unscreened programs use rules that give priority to those who attend information sessions or open houses. Screened programs as well as Screened: Language programs select students by individually assorted measure such as elementary school GPA, statewide test scores, punctuality and interviews.

[49]In the current data set, the finest level of geographic information of a student is census-tract level. The distance between students and schools is calculated as follows. For each census tract in New York City, we use the latitude and longitude coordinates of the centroid from corresponding year's US Census gazetteer file. School's coordinates are calculated using their exact street addresses with Google API. Next we calculate the distance between the coordinates of the exact school location and students' census tract of residence centorid based on the Haversine formula.

Zoned programs guarantee admissions or give priority to students who reside in the school's zone, and Talent Test programs use auditions as the main criteria.

High school programs use similar admission methods as middle schools—Unscreened, Limited Unscreened, Screened, Screened: Language, Screened: Language & Academics, Zoned, Audition, Educational Option, and Continuing 8th Graders. Audition programs are similar to Talent Test middle school programs, and Educational Option is a mixture of Unscreened and Screened. Educational Option programs have the purpose of serving students at diverse academic performance levels. These programs divide students into high (highest 16%), middle (68%) and low ELA (lowest 16%) levels. 50% of the seats in each group are filled using school-specific criteria like a screened program and the other 50% are filled randomly similarly as an unscreened program. (*NYC DOE Introduction to High School Admissions*) Continuing 8th Graders programs are open only to continuing eighth graders in the same school. Other admission methods are similar to middle school choice.

Sample Restriction We start with 72,318 observations in the middle school application data. Out of 72,318, 67,153 students participated in the main round of the middle school application. We drop students with missing demographic characteristics or invalid standardized test scores, and are left with 62,972 students. Among the remaining students, 54,012 students participated in high school application after three years.[50] We present summary statistics and balance test results of these 54,012 students in section 3.2.[51] For new middle and high schools, school characteristics are missing. After excluding students who went to a new middle school and whose high school ranked ordered list is filled only with new high schools, we have 44,237 students. The estimates in Table 2.5 are derived based on this sample.

[50]Those who participated in the middle school choice but not participated in the high school choice do not appear in the data afterwards. Examples might include drop-outs, those who attend private or charter high schools, and those who moved out of NYC. These are more likely to be low performers, subsidized lunch status, or Black students.

[51]801 students applied only to new middle schools on which there is no characteristics of the previous cohort. We present summary statistics and balance test results on middle school application behavior for the rest (=53,211).

2.6.2 Student-Proposing Deferred Acceptance Algorithm

In detail, SPDA works as follow:

- **Step 1**
  Each student proposes to her first choice. Each program tentatively assigns seats to its proposers one at a time, following their priority order. The student is rejected if no seats are available at the time of consideration.

- **Step $k \geq 2$**
  Each student who was rejected in the previous step proposes to her next best choice. Each program considers the students it has tentatively assigned together with its new proposers and tentatively assigns its seats to these students one at a time following the program's priority order. The student is rejected if no seats are available when she is considered.

- The algorithm terminates either when there are no new proposals or equally when all rejected students have exhausted their preference lists.

SPDA produces the student-optimal stable matching and is strategyproof i.e., truth-telling is a dominant strategy for students.

2.6.3 An Example of Calculating Propensity Scores

The following example illustrates how to calculate the propensity score.[52]

Consider student i who submits a rank-ordered list A-B-C where A is her most preferred option and C is her least preferred option. Priority score used for admissions is a sum of priority group and a tie-breaker, where priority group lexico-

[52]Note that the propensity score in this context denotes the exact probability of being treated, and involves no prediction of the odds by estimating a logit or a probit model, which is typically found in papers with propensity score matching (for example, Dehejia and Wahba, 2002; Smith and Todd, 2005).

graphically dominates tie-breakers. That is, student i's score at program j is

$$\text{score}_{ij} = \underbrace{PG_{ij}}_{\text{priority group}\in\mathbb{N}} + \underbrace{TB_{ij}}_{\text{tie-breaker}\in[0,1]}$$

where i has higher priority than i' at j if and only if $\text{score}_{ij} > \text{score}_{i'j}$. Programs A and B share a random tie-breaker $TB_{iA} = TB_{iB} \overset{iid}{\sim} U[0,1]$, and programs C uses a non-random tie-breaker $TB_{iC} \sim F_i$, where F_i is unknown and potentially depends on the student and has a support on [0,1]. A cutoff of program j is given by the minimum of scores of admitted students at j if all seats are filled, and $-\infty$ if some seats are left unfilled. Assuming a large market (Azevedo and Leshno, 2016b; Fack et al., 2019b; Calsamiglia et al., 2020), student i is admitted to program j if $\text{score}_{ij} > \text{cutoff}_j$ and at the same time rejected from all programs ranked above j.

Programs	A	B	C	D
PG_{ij}	1	1	2	2
Cutoff	2.2	1.4	2.6	$-\infty$
Admission Prob.	0	1×0.6	$1 \times 0.4 \times (1 - F_i(0.6))$	$1 \times 0.4 \times F_i(0.6)$
Local Admission Prob.	0	1×0.6	$1 \times 0.4 \times \mathbf{0.5}$	$1 \times 0.4 \times \mathbf{0.5}$

Table 2.11: Example of Propensity Score

Table 2.11 illustrates how to calculate the propensity score for student i in this example. Student i has no chance of being admitted to program A, since no realization of the tie-breaker is large enough to clear the cutoff of program A. Next, the probability of being assigned to program B is the probability of being rejected from program A (=1) times the probability of getting accepted to program B. The cutoff of B is 1.4, so i can be assigned to program B as long as her lottery number is greater than 0.4, which happens with a probability of 0.6. Hence student i's admission probability at program B is $1 \times 0.6 = 0.6$. Next, i gets assigned to program C if she is rejected from all previous options (=1×0.4) and then clears the cutoff of program C. While it is impossible to get the exact probability of clearing the cutoff, $F_i(0.6)$, Theorem 1 of Abdulkadiroğlu et al. (2021) suggests that i clears the cutoff with half chance if i's tie-breaker TB_{iC} is close enough to the cutoff.

2.6.4 Additional Procedures for Computation

Constructing Priority Scores and Simulating Uncertainties It is necessary to construct each student's priority to use ex-post stability. First, to calculate the continuation value, we need to simulate the set of feasible schools in each realization of ex-post cutoffs by running SPDA algorithm multiple times, which takes students' priorities when attending different middle school programs. Second, to interpret data as a conditional multinomial logit model, we need to construct the feasible set of programs to the student, regardless of if she ranked them or not.

Priority scores consist of mainly three ingredients: eligibility, priority group, and priority ranks at programs involving screening. We provide details on how we construct each of the ingredients.

First, eligibility and priority group are determined in a deterministic manner, based on the pre-announced rule in NYC Middle School Directory and NYC High School Directory published every year before public school applications.

Next, we estimate the priority ranks for Screened, Screened: Language, Screened: Language & Academics, and the screened part of Education Option programs. While the data set includes the priority rank of applicants to each program, there is no information on the ranks of those who did not apply to that particular program. In addition, the exact formula that each program uses is not known to us. To this end, we assume there exists a program-specific latent variable as a function of student characteristics, which determines the rank of students at each program. Specifically, let w_{ij} be the latent variable of i at actively ranking program j, as a function of student characteristics Z_i. We assume:

$$w_{ij} = \beta_j Z_i + e_{ij}$$

and

$$i \text{ is ranked higher than } i' \text{ if and only if } w_{ij} > w_{i'j}$$

where Z_i includes standardized statewide Math and ELA exam scores; Math, Social Sciences, English, Science GPA; and days absent and days late. We assume e_{ij} is iid as EVT1. From the data, we gather all possible pairs of applicants to program j, and maximize the following likelihood:

$$\hat{\beta}_j = \text{argmax}_b \sum_{i>i',i,i'\in \mathcal{I}_j} \log \frac{1}{\exp(w_{ij}) + \exp(w_{i'j})}$$
$$\cdot \left(\exp(w_{ij}) 1\{i \text{ is ranked higher than } i'\} + \exp(w_{i'j}) 1\{i \text{ is ranked lower than } i'\}\right)$$

Using the estimates $\hat{\beta}_j$, we predict $\hat{w}_{ij} = \hat{\beta}_j Z_i$ for all i and reconstruct the priority ranks based on $\hat{w}_{ij}$.

Finally, we describe how to simulate the uncertainties in cutoffs, ω. To take care of two sources of uncertainties in the cutoffs the student is facing – uncertainty in other student's types and lottery draws – we bootstrap 200 times from the data and create multiple economies, draw 200 sets of lotteries and run SPDA algorithm $200 \times 200 = 40,000$ times. We use the resulting empirical distribution of cutoffs as the distribution of cutoffs.

Evolution of Test Scores Consider any time-variant student characteristics y_i^M and y_i^H that are part of Z_i^M and Z_i^H. That is, they are the same type of variable but may change as a function of middle school attendance when the student applies to middle schools and high schools. For example, the test scores of a student may change depending on the middle school she attends, because different middle schools may have different effectiveness. We estimate each middle school's 'production function'[53] using a value-added model.

[53]To ensure enough sample size, we estimate the value-added of each middle school instead of middle school program.

Specifically, let $y^H_{i,m}$ be the potential y^H_i when student i attends middle school m. Then based on the 'selection on observables' assumption,

$$\mathsf{E}\left[y^H_{i,m} \mid Z^M_i, m\right] = \alpha_m + Z^{M\prime}_i \beta_m, \quad m \in \mathcal{M}$$

We estimate via OLS of $y^H_{i,m(i)}$ on school indicators interacted with Z^M_i where $m(i)$ is the actual middle school attendance in the data.

Table 2.12 reports the mean and standard deviations of the coefficients $\hat{\alpha}_m, \hat{\beta}_m$ and their standard errors across middle schools. First, students with higher baseline test scores tend to have higher test scores, reflecting their higher academic ability. Second, there exist significant variation across schools as well as heterogeneity based on student observable characteristics.

	Math		ELA	
	Coefficient	SE	Coefficient	SE
Baseline Test Score	0.346	0.035	0.331	0.033
	(0.060)	(0.015)	(0.040)	(0.013)
Female	1.591	1.650	3.077	1.517
	(1.425)	(0.412)	(2.327)	(0.352)
Asian	6.002	3.993	6.029	3.402
	(4.892)	(2.108)	(4.617)	(1.547)
Black	-2.422	4.542	-2.502	4.642
	(6.194)	(2.527)	(3.826)	(3.216)
Hispanic	-2.309	2.708	-0.738	2.472
	(3.945)	(1.260)	(3.391)	(1.008)
English Language Learner	-2.862	5.691	1.239	6.066
	(7.230)	(2.669)	(6.273)	(3.045)
Student with Disability	-6.885	2.345	-5.571	2.212
	(3.192)	(0.690)	(2.122)	(0.663)
Free or Subsidized Lunch	-1.380	2.264	-1.501	2.013
	(2.124)	(1.190)	(1.974)	(0.863)

Table 2.12: Mean and Standard Deviation of VA Coefficients Across Schools

Computation of Continuation Value By the assumptions on unobservables and ex-post stability, the continuation value term in Equation 2.9 can be simplified to

$$
\begin{aligned}
& E_{\gamma_i^H,\omega,\eta_i,Z_i^H}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|Z_i^M,\gamma_i^M,\epsilon_i,m\right] \\
=& \int_\omega E_{\xi_i,\eta_i}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|Z_i^M,\gamma_i^M,m\right] dH(\omega|Z_i^M,\gamma_i^M,\epsilon_i,m) \\
=& \int_\omega\int_{\xi_i} E_{\eta_i}\left[\max_{j\in O_i(Z_i^H,m;\omega)} V_{ij}\middle|Z_i^M,\gamma_i^M,m,\xi_i\right]\phi(\xi_i|\Sigma_\xi)d\xi_i dH(\omega|Z_i^M,\gamma_i^M,\epsilon_i,m) \\
=& \int_\omega\int_{\xi_i}\left(\mu+\log\left(\sum_{j\in O_i(Z_i^H,m;\omega)}\exp(\nu_{ij}(\xi_i))\right)\right)\phi(\xi_i|\Sigma_\xi)d\xi_i dH(\omega|Z_i^M,\gamma_i^M,\epsilon_i,m)
\end{aligned}
$$

where μ is the Euler-Mascheroni constant. The first integral over ω is calculated by using the empirical distribution of ω as described in section 2.6.4. The second integral over ξ_i is calculated using sparse grids quadratures (Heiss and Winschel, 2008). We use students' residence in the first period to calculate the distance to each high school program in the calculation of the continuation value.

Likelihood Function Let student i's assigned middle and high school programs be m_i, j_i and the respective feasible sets be O_i^m, O_i^h. Let u_{im} and ν_{ij} denote the part of U_{im} and V_{ij} excluding the idiosyncratic preference terms ϵ_{im} and η_{ij}. Also, denote the parameters to be estimated as

$$
\theta = \left(\beta_0^M, \beta_Z^M, \beta_0^H, \beta_Z^H, \rho_0, \{\rho_\tau\}_\tau, \Sigma_\gamma, \Sigma_\xi, \vartheta_m, \vartheta_h, \delta, \lambda^M, \lambda^H\right)
$$

Then for student i, conditional on γ_i^M, ξ_i,

$$\begin{aligned} P_i(\theta, \gamma_i^M, \xi_i) &= P(\text{observe } m_i, j_i | \gamma_i^M, \xi_i, \theta) \\ &= P\left(\begin{array}{cc} U_{im_i} = \max_{m \in O_i^m} U_{im} & \text{and} \\ V_{ij_i} = \max_{j \in O_i^h} V_{ij} \text{ given } m_i \end{array} \middle| \gamma_i^M, \xi_i, \theta \right) \\ &= \frac{\exp(u_{im_i}(\gamma_i^M, \theta))}{\sum_{m \in O_i^m} \exp(u_{im}(\gamma_i^M, \theta))} \frac{\exp(v_{ij_i}(\gamma_i^M, \xi_i, \theta; m_i))}{\sum_{j \in O_i^h} \exp(v_{ij}(\gamma_i^M, \xi_i, \theta; m_i))} \end{aligned}$$

where the second equality comes from the ex-post stability and the third equality comes from the distributional assumptions on the unobservables. Then,

$$P_i(\theta) = \int_{\gamma_i^M} \int_{\xi_i} P_i(\theta, \gamma_i^M, \xi_i) \phi(\xi_i | \Sigma_\xi) \phi(\gamma_i^M | \Sigma_\gamma) d\xi_i d\gamma_i^M$$

where $\phi(\cdot|\Sigma)$ is the pdf of a multivariate normal with zero mean and covariance matrix Σ, and hence

$$\prod_i P_i(\theta), \text{ or } \sum_i \log P_i(\theta)$$

is the final likelihood function to be maximized.

2.6.5 Alternative Specifications

Static Model The static model has the same main components of the main model, but with three marked differences. First, we assume students are myopic so that they do not take the high school application into consideration when they make middle school choices ($\delta = 0$). Second, we do not allow the unobserved tastes on program characteristics to be serially correlated ($\rho_0 = 0$), and third, middle school type effects are absent ($\rho_\tau = 0, \forall \tau$). Table 2.13 reports the preference estimates of the static model. The goodness-of-fit measures are reported together with our main specification in Table 2.16.

	Middle Schools			High Schools		
	est	se		est	se	
Panel A: Preference Estimates						
Fraction of High Performer						
Main Effect	-7.574	(1.273)	***	0.929	(0.287)	***
Asian	-0.788	(1.773)		0.870	(0.411)	**
Black	10.573	(1.915)	***	-0.169	(0.469)	
Hispanic	2.682	(1.246)	**	-0.356	(0.339)	
Free or Reduced Lunch	0.804	(1.039)		-1.008	(0.280)	***
English Language Learner	-0.988	(2.358)		0.523	(1.227)	
5th Grade Test Score	0.881	(0.545)		1.995	(0.165)	***
Fraction of White						
Main Effect	8.805	(1.382)	***	6.264	(0.401)	***
Asian	0.621	(1.458)		-1.513	(0.642)	**
Black	-10.161	(1.828)	***	-2.347	(0.645)	***
Hispanic	-2.729	(1.090)	**	-1.654	(0.447)	***
Free or Reduced Lunch	-1.628	(0.841)	*	-0.231	(0.371)	
English Language Learner	-0.009	(2.030)		-1.288	(1.238)	
5th Grade Test Score	-0.597	(0.435)		-0.109	(0.194)	
1(STEM)						
Main Effect	0.396	(0.260)		-0.679	(0.124)	***
Asian	-0.076	(0.273)		-0.144	(0.203)	
Black	-0.690	(0.294)	**	0.078	(0.192)	
Hispanic	-0.206	(0.207)		0.089	(0.142)	
Free or Reduced Lunch	-0.135	(0.176)		0.241	(0.127)	*
English Language Learner	0.208	(0.298)		0.874	(0.321)	***
5th Grade Test Score	0.113	(0.081)		-0.030	(0.063)	
Panel B: Middle School Type Effects						
Type 1 (High Achievement MS)						
Fraction of High Performer				0.159	(0.288)	
Fraction of White				0.627	(0.391)	
1(STEM)				-0.307	(0.137)	**
Type 2 (High Minority MS)						
Fraction of High Performer				0.936	(0.321)	***
Fraction of White				-2.770	(0.479)	***
1(STEM)				0.173	(0.135)	
Panel C: Unobservable Tastes						
(1,1) of Variance of Random Taste	40.970	(13.423)	***	0.730	(0.408)	*
(1,2)	-37.851	(12.154)	***	-3.058	(1.052)	***
(1,3)	-1.330	(2.585)		0.496	(0.164)	***
(2,2)	35.893	(13.309)	***	12.813	(2.760)	***
(2,3)	1.790	(3.262)		-2.078	(0.492)	***
(3,3)	0.384	(0.721)		0.337	(0.159)	**
Panel D: Other Parameters						
Outside option	-1.851	(0.198)	***	-0.112	(0.173)	
Distance	0.718	(0.031)	***	0.496	(0.018)	***

Notes: We report the preference estimates of the static model. School characteristics 'Fraction of High Performer' and 'Fraction of White' are between 0 and 1, and '1(STEM)' is an indicator variable. In Panel A, Main Effect is the common taste (β_0^M, β_0^H), and we also include interactions of each school characteristics with Asian, Black, Hispanic, Free or Reduced Lunch (FRL) status, English Language Learner (ELL) status, 5th Grade Test Score in the following rows (β_Z^M, β_Z^H). Robust standard errors are reported in parentheses.

Strict Truth-Telling (STT) We also estimate the model with a different assumption on student behavior. Strict Truth-Telling (STT) assumes that (Fack et al., 2019b)

1. Students rank all acceptable programs (i.e., better than the outside option) in their true preference order.

2. All unranked programs are unacceptable to the student. That is, they are worse than the outside option.

Hence, the likelihood used in the estimation for STT is as follows.

Let student i's ROL on middle school programs and high school programs be $R_i^M = (R_{i,1}^M, \cdots, R_{i,|R_i^M|}^M)$ and $R_i^H = (R_{i,1}^H, \cdots, R_{i,|R_i^H|}^H)$ respectively. We will use the notation that $c \succ_i c'$ to denote that programs c and c' are both ranked by student i and c is ranked higher than c' in her ROL. Note that $\{c' \in R_i : c' \not\succ_i c\}$ includes c itself.

Then for student i, conditional on γ_i^M, ξ_i,

$$
\begin{aligned}
P_i(\theta, \gamma_i^M, \xi_i) =& P(\text{observe } R_i^M, R_i^H | \gamma_i^M, \xi_i, \theta) \\
=& P\left(\begin{array}{c} U_{i,R_{i,1}^M} > \cdots > U_{i,R_{i,|R_i^M|}^M} > U_{i0_m} > U_{im'}, \forall m' \in \mathcal{M} \setminus R_i^M \text{ and} \\ V_{i,R_{i,1}^H} > \cdots > V_{i,R_{i,|R_i^H|}^H} > V_{i0_h} > V_{ij'}, \forall j' \in \mathcal{J} \setminus R_i^H \end{array} \middle| \gamma_i^M, \xi_i, \theta \right) \\
=& \frac{\exp(u_{i0_m})}{\exp(u_{i0_m}) + \sum_{m' \notin R_i^M} \exp(u_{im'}(\gamma_i^M, \theta))} \\
& \times \prod_{m \in R_i^M} \left(\frac{\exp(u_{im}(\gamma_i^M, \theta))}{\exp(u_{i0_m}) + \sum_{m' \not\succ_i m} \exp(u_{im'}(\gamma_i^M, \theta))} \right) \\
& \times \frac{\exp(v_{i0_h})}{\exp(v_{i0_h}) + \sum_{j' \notin R_i^H} \exp(v_{ij'}(\gamma_i^H, \xi_i, \theta))} \\
& \times \prod_{j \in R_i^H} \left(\frac{\exp(v_{ij}(\gamma_i^H, \xi_i, \theta))}{\exp(v_{i0_h}) + \sum_{j' \not\succ_i j} \exp(v_{ij'}(\gamma_i^H, \xi_i, \theta))} \right)
\end{aligned}
$$

Then,

$$P_i(\theta) = \int_{\gamma_i^M} \int_{\xi_i} P_i(\theta, \gamma_i^M, \xi_i) \phi(\xi_i | \Sigma_\xi) \phi(\gamma_i^M | \Sigma_\gamma) d\xi_i d\gamma_i^M$$

where $\phi(\cdot|\Sigma)$ is the pdf of a multivariate normal with mean zero and covariance matrix Σ, and hence

$$\prod_i P_i(\theta), \text{ or } \sum_i \log P_i(\theta)$$

is the likelihood.

Table 2.15 reports the preference estimates from the alternative model using STT. The main difference we find is the negative and statistically significant estimate on δ, the discount factor. The intuition is as follows. In case of payoff-irrelevant mistakes in which students omit favorable yet infeasible middle school programs from their rank-ordered list, STT interprets those omitted programs as less preferred than all ranked programs as well as the outside option. However, these competitive programs with high cutoffs will have high continuation value since they are likely to provide higher opportunities of getting into favorable high school programs. As a result, STT would incorrectly infer that middle schools with high continuation value as unfavorable, resulting in a negative estimate for the discount factor.

The goodness-of-fit measures are reported in Table 2.14. STT overall is outperformed by ex-post stability. Especially, the mean predicted fraction of students assigned to the observed assignments is nearly decreased to half (see Table 2.16), reconfirming the fact that truth-telling assumption may be problematic even in a strategyproof environment.

2.6.6 Additional Goodness-of-Fit Measures

We provide additional goodness-of-fit measures along two dimensions: how well it predicts the assignments, and how well it predicts students' revealed preferences.

First, Panel A of Table 2.16 compares each student's predicted assignment to

	Dynamic, STT	
	MSAP	HSAP
Panel A. Simulated versus observed assignment (100 simulated samples)		
Mean predicted fraction of students	0.3129	0.1111
assigned to observed assignments	(0.0050)	(0.0048)
Panel B. Predicted versus observed partial preference order		
Mean predicted probability that a student's partial preference order among the programs in her ROL coincides with the submitted rank order	0.3769	0.1261

Notes: Panel A calculates the average success rate of predicting the observed assignments in the data using the model estimate. 100 sets of unobservable variables ($\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij}$) are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and SPDA are run, resulting in 5,000 simulated assignments. The average and standard deviation (in parentheses) across the unobservable draws are reported. Panel B calculates the average predicted probability of a student's partial preference order among the programs ranked on her ROL coincides with the submitted rank order.

Table 2.14: Strict Truth-Telling: Goodness of Fit

the observed one. We have about 57.1% success rate for middle schools and 20.2% for high schools.[54] The higher success rate for middle schools can be explained by the fact that first, the number of programs are much smaller for middle schools, and second, middle schools have much more 'Zoned Guarantee' programs that guarantee admissions to students who are zoned to the school as long as they rank them.

Next in Panel B, we take as given the programs that a student has included in her submitted ROL, and compute the probability of observing this particular preference order among the ranked programs. Given the distributional assumptions

[54]We provide two benchmarks to evaluate if those measures are good or bad. First, the upper bound is calculated using submitted ROLs in the data, without relying on any estimates or the model: 78.0% for MSAP and 61.2% for HSAP. They do not equal 100% due to lottery draws (8 (40%) middle school programs and 29 (62%) high school programs use lottery draws for tie-breaking in our sample). Next, the lower bound is calculated using random prediction. That is, we let students randomly apply to programs and programs randomly select who to admit. On average, we find 5.9% for MSAP and 2.3% for HSAP.

	Middle Schools			High Schools		
	est	se		est	se	
Panel A: Preference Estimates						
Fraction of High Performer						
Main Effect	4.385	(0.335)	***	-0.628	(0.141)	***
Asian	0.87	(0.527)	*	0.708	(0.209)	***
Black	1.253	(0.460)	***	0.601	(0.231)	***
Hispanic	-0.042	(0.376)		-0.278	(0.164)	*
Free or Reduced Lunch	0.080	(0.324)		-0.217	(0.133)	
English Language Learner	-0.218	(0.806)		0.721	(0.481)	
5th Grade Test Score	0.959	(0.172)	***	1.907	(0.066)	***
Fraction of White						
Main Effect	0.496	(0.190)	***	3.659	(0.139)	***
Asian	0.328	(0.354)		-0.693	(0.226)	***
Black	0.417	(0.327)		-1.102	(0.242)	***
Hispanic	0.605	(0.256)	**	-0.239	(0.163)	
Free or Reduced Lunch	0.170	(0.217)		-0.202	(0.132)	
English Language Learner	-0.070	(0.565)		-1.205	(0.455)	***
5th Grade Test Score	-0.414	(0.114)	***	-0.904	(0.062)	***
1(STEM)						
Main Effect	0.630	(0.069)	***	-0.305	(0.065)	***
Asian	-0.241	(0.129)	*	0.201	(0.102)	**
Black	-0.468	(0.113)	***	-0.051	(0.102)	
Hispanic	-0.200	(0.091)	**	0.122	(0.074)	*
Free or Reduced Lunch	-0.205	(0.080)	**	-0.010	(0.064)	
English Language Learner	0.175	(0.169)		0.286	(0.175)	
5th Grade Test Score	0.074	(0.041)	*	-0.264	(0.026)	***
Panel B: Middle School Type Effects						
Type 1 (High Achievement MS)						
Fraction of High Performer				0.290	(0.141)	**
Fraction of White				0.321	(0.134)	**
1(STEM)				-0.101	(0.067)	
Type 2 (High Minority MS)						
Fraction of High Performer				1.320	(0.177)	***
Fraction of White				-0.774	(0.177)	***
1(STEM)				0.024	(0.076)	
Panel C: Unobservable Tastes						
ρ_0				0.457	(0.095)	***
				0.495	(0.279)	*
				0.147	(0.097)	
(1,1) of Σ_γ	3.307	(0.841)	***			
(1,2)	-0.930	(0.373)	**			
(1,3)	-1.237	(0.218)	***			
(2,2)	0.262	(0.148)	*			
(2,3)	0.348	(0.108)	***			
(3,3)	0.463	(0.093)	***			
(1,1) of Σ_ξ				3.040	(0.306)	***
(1,2)				-2.538	(0.250)	***
(1,3)				-0.344	(0.092)	***
(2,2)				2.312	(0.265)	***
(2,3)				-0.009	(0.078)	
(3,3)				0.491	(0.054)	***
Panel D: Other Parameters						
Outside option	-0.528	(0.256)	**	1.974	(0.038)	***
Distance	0.746	(0.012)	***	0.486	(0.007)	***
Discount Factor	-0.65	(0.062)	***			

Notes: We report the preference estimates of the model based on STT. School characteristics 'Fraction of High Performer' and 'Fraction of White' are between 0 and 1, and '1(STEM)' is an indicator variable. In Panel A, Main Effect is the common taste (β_0^M, β_0^H), and we also include interactions of each school characteristics with Asian, Black, Hispanic, Free or Reduced Lunch (FRL) status, English Language Learner (ELL) status, 5th Grade Test Score in the following rows (β_Z^M, β_Z^H). Robust standard errors are reported in parentheses.

	Dynamic Model		Static Model	
	MSAP	HSAP	MSAP	HSAP
Panel A. Simulated versus observed assignment (100 simulated samples)				
Mean predicted fraction of students assigned to observed assignments	0.5709 (0.0053)	0.2022 (0.0049)	0.5539 (0.0058)	0.2018 (0.0048)
Panel B. Predicted versus observed partial preference order				
Mean predicted probability that a student's partial preference order among the programs in her ROL coincides with the submitted rank order	0.3848	0.1395	0.3215	0.1422
Panel C. Likelihood Ratio Test				
H_0: $\delta = \rho_0 = \rho_\tau = 0, \forall\tau$			Reject H_0 ($p < 0.001$)	

Notes: Panel A calculates the average success rate of predicting the observed assignments in the data using the model estimate. 100 sets of unobservable variables ($\gamma_i^M, \xi_i, \epsilon_{im}, \eta_{ij}$) are drawn and for each set, 50 sets of tie-breaking lotteries are drawn and SPDA are run, resulting in 5,000 simulated assignments. The average and standard deviation (in parentheses) across the unobservable draws are reported. Panel B calculates the average predicted probability of a student's partial preference order among the programs ranked on her ROL coincides with the submitted rank order.

Table 2.16: Additional Goodness of Fit Measures

on $\gamma_i^M, \xi_i^H, \epsilon_{im}, \eta_{ij}$, we can calculate the probabilities without relying on Monte Carlo simulations. We have 38.5% for middle schools and 14.0% for high schools. The difference in probabilities between middle schools and high schools is due to the larger number of high school programs and longer high school ROLs.[55]

Recall the key features of the dynamic model: forward-looking agents, serial correlation of the unobservable tastes, and middle school type effects. To highlight the importance of including those features in the model, we estimate a restricted static model without the dynamic components of the model (i.e., no forward-looking, no serial correlation of unobserved tastes, no middle school type effects).[56]

The goodness-of-fit measure of the restricted model is worse than the full dynamic model in terms of middle school applications, and more or less similar in

[55]On average, students rank 2 middle school programs (std 1.22), and 4 high school programs (std 2.60).

[56]See subsection 2.6.5 for a more detailed description and results on preference estimates.

terms of high school applications. This is as expected as the static model does not consider the forward-looking behavior of students in the middle school application stage, and thus does a worse job of fitting the corresponding data. On the other hand, high school application is the *last* stage of the multi-period game, and hence unlikely to be affected by whether including a dynamic feature or not.

Since the static model is a nested model of the full dynamic model in which the restriction that ρ_0, δ and $\rho_\tau, \forall\tau$ are equal to zero is imposed, we can perform a likelihood ratio (LR) test. The result is reported in Panel C. The static model is strongly rejected in favor of our main dynamic model ($p < 0.001$), reconfirming the importance of including the forward-looking behavior of students, middle school effects on tastes, and serial correlation of unobservable tastes in the model.

2.6.7 Segregation Measures

Theil's H Index Theil's H Index is also known as the Information Theory Index or the Multigroup Entropy Index. In this paper, we closely follow the definition used by the United States Census Bureau to describe housing patterns (Iceland, 2004).[57]

First, the entropy score of the entire economy is calculated as:

$$E = \sum_{r=1}^{R} (\Pi_r) \log(1/\Pi_r)$$

where Π_r is a particular racial group r's proportion in the whole population in the economy. The entropy score measures the diversity in the economy, where higher number indicates higher diversity.

Next, for each school $j = 1, 2, \cdots, J$, the entropy score of j is calculated similarly:

$$E_j = \sum_{r=1}^{R} (\Pi_{r,j}) \log(1/\Pi_{r,j})$$

[57]See https://www.census.gov/topics/housing/housing-patterns/about/multi-group-entropy-index.html

where $\Pi_{r,j}$ is a racial group r's proportion in the whole population in school j.

Finally, Theil's H index is calculated as the weighted average of deviation of each j's entropy from the entropy score of the entire economy, where the weight is the number of students at each school:

$$H = \sum_{j=1}^{J} \left[\frac{t_j(E - E_j)}{E \cdot T} \right]$$

where t_j is the total number of students in school j, and $T = \sum_{j=1}^{J} t_j$ is the total number of students in the economy. By construction, H is between 0 and 1 where 0 means maximum integration (i.e., all schools have the same racial composition as the whole economy), and 1 means maximum segregation.

Sorting Index Sorting index for a given characteristic is defined by the ratio of the between-school variance to the total variance, measuring the fraction of the variance in a given characteristic that can be explained by the between-school differences. Specifically, let y_{ij} be the student i's characteristic of interest who is enrolled in j. Then, the sorting index for y is simply obtained by the R^2 of the following linear regression:

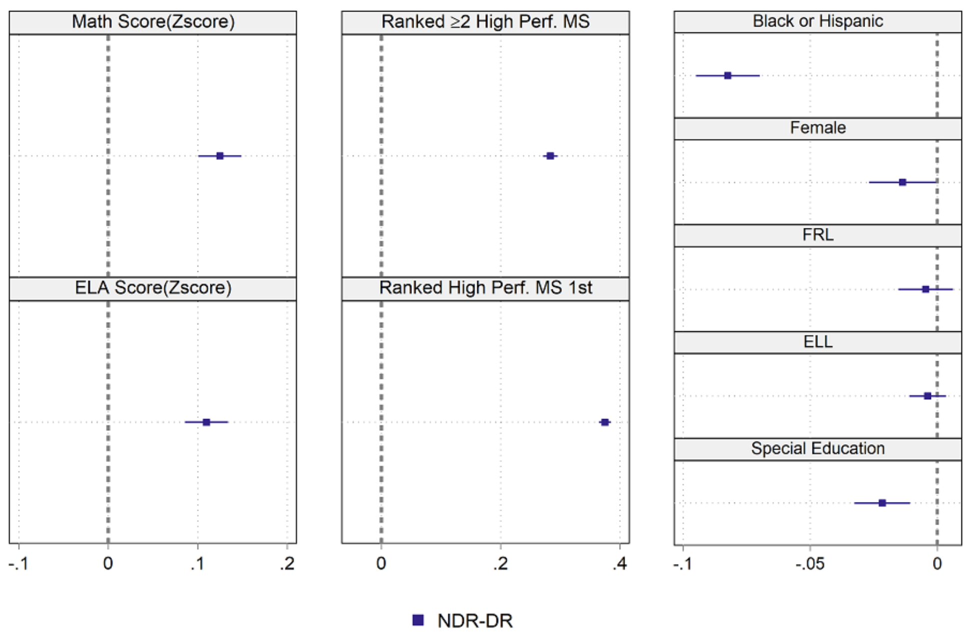
$$y_{ij} = \alpha_j + e_{ij}$$

It varies between 0 and 1 by definition, and 0 means maximum integration, and 1 means maximum segregation.

2.6.8 Additional Tables and Figures

Additional Tables and Figures from section 2.3

Balance Test: Nondegenerate Risk versus Degenerate Risk Sample Figure 2.5 presents the mean difference between those with non-degenerate offer risk and degenerate (0 or 1) offer risk, when the treatment variable is 'attended a high achievement middle school'. In our data, 2/3 of degenerate risk sample have the propensity score equal to 0, which means they did not apply to any of the high achievement middle schools or had zero chance of getting in conditional on applying, suggesting they are different from the non-degenerate risk sample. Indeed, we find that students with non-degenerate risk and those with degenerate risk are quite different: students with non-degenerate risk have higher test scores, and less likely to be Black or Hispanic, and obviously ranked many treatment middle schools. It reconfirms that the 2SLS estimates we find in subsection 2.3.2 are local average treatment effect (LATE).



Notes: This table shows the t-test results of covariate mean difference between those with non-degenerate offer risk and those with degenerate offer risk. Markers show the exact estimates, and 95% CIs are presented. Robust standard errors are estimated (N=50,871).

Figure 2.5: Covariate Balance Test: Nondegenerate v.s. Degenerate Risk Samples

Robustness Check We further investigate if the reduced-form evidence of middle school effects on high school choice we identified is mainly driven by the increase in students' test scores. Moreover, we control for the length of high school application list because students submit lists of different lengths, and the average characteristics of schools change along the rank on the list, as shown in Table 2.3, Table 2.4.

First, Table 2.17 uses the same identification strategy as in section 2.3 to show that attending a high achievement middle school has a causal impact on the increase of students' 8th grade math test scores and shorter high school ROLs.

Nevertheless, Table 2.18 shows that the 2SLS estimates are robust to controlling for the middle school endline test score and the length of ROLs. Notably, students with higher endline test scores apply high schools with better academic performance, the pattern described well in previous studies that estimate school demand (e.g., Hastings et al., 2005; Abdulkadiroğlu et al., 2017a). For example, the estimates in Column (2) illustrate that the increase of student's ELA test score by 1 standard deviation is associated with 0.9 pp increase in the average graduation rate of high schools on her application list. With the test score controlled, the effect of attending a high achievement middle school is 1.6 pp, which is comparable to the effect in Column (4) of Table 2.5. This mediation analysis shows that there is still a treatment effect of attending a high achievement middle school, even controlling for test scores. Motivated by this finding, we include a separate component that captures the effect of middle school beyond its value-added on test scores in the structural model presented in section 2.4.

Heterogeneous Treatment Effects Next, we provide results on possible heterogeneous treatment effects. Figure 2.6 presents the heterogeneity analysis results by student demographic characteristics. The effect is broad-based, but there is a larger effect among students whose baseline mean of the outcome variable is lower than their peers. For example, while attending a high achievement middle school increases the college enrollment rate of an ELL student's assigned school by 10 pp, it does by 3.1 pp for a non-ELL student. The baseline level of the matched school's college enrollment rate is 59 and 66 percent among ELL/non-ELL students,

	(1)	(2)	(3)
Model	OLS	2SLS	2SLS
Sample	All	All	NDR
Panel A: 8th Grade ELA Score (Zscore)			
From High Achievement MS	0.064***	0.032	0.032
	(0.019)	(0.035)	(0.042)
N	42516	42516	6826
R2	0.610	0.616	0.639
$\bar{y}$	0.090	0.090	0.183
Panel B: 8th Grade Math Score (Zscore)			
From High Achievement MS	0.173***	0.115**	0.152**
	(0.029)	(0.056)	(0.067)
N	32935	32935	5562
R2	0.582	0.591	0.651
$\bar{y}$	0.051	0.051	0.202
Panel C: Length of Application List			
From High Achievement MS	-0.718***	-0.716*	-1.161**
	(0.265)	(0.429)	(0.585)
N	44237	44237	7062
R2	0.117	0.171	0.360
$\bar{y}$	7.579	7.579	7.049
Panel D: 1(Assigned to the First Ranked School)			
From High Achievement MS	0.006	0.062*	0.035
	(0.015)	(0.033)	(0.042)
N	41312	41312	6571
R2	0.037	0.053	0.154
$\bar{y}$	0.472	0.472	0.464

Notes: Standard errors clustered at graduating middle school in parentheses, *p<0.1, **p<0.05, ***p<0.01. All regressions control for student ethnicity, gender, English Language Learner status, Free-Reduced Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. All columns also control for saturated dummy for all possible values of propensity score of being assigned to a high achievement MS, and local linear control for non-random tie-breakers.

Table 2.17: Effect of Attending a High Achievement MS on Other Outcomes

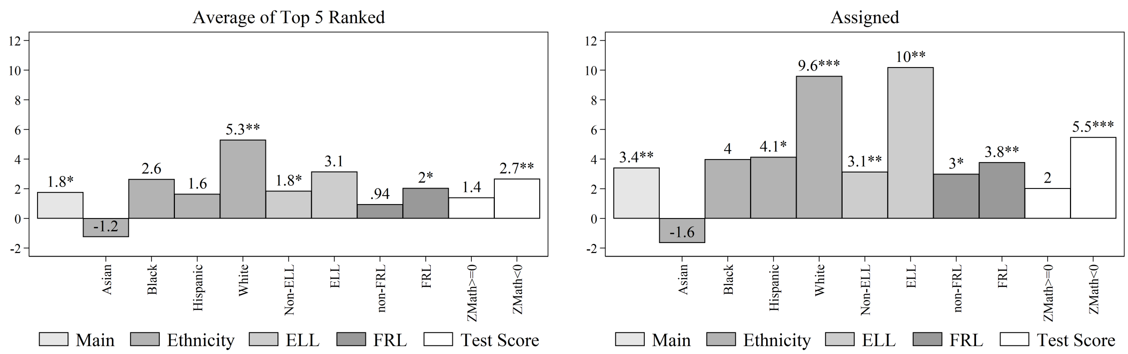
Model	(1) 2SLS	(2) 2SLS	(3) 2SLS	(4) 2SLS
Dependent Variable	Average of Top 5		Matched	
Sample	NDR	NDR	NDR	NDR
Panel A: 4yr Graduation Rate (%)				
From High Achievement MS	1.353*	1.565**	2.341**	2.274**
	(0.700)	(0.766)	(1.119)	(1.104)
8th Grade ELA Score	0.909***		1.601***	
	(0.170)		(0.288)	
8th Grade Math Score	0.765***		1.006***	
	(0.206)		(0.326)	
Number of Programs Ranked		0.160**		-0.126
		(0.072)		(0.086)
N	7060	7060	6687	6687
R2	0.398	0.390	0.264	0.253
ȳ	83.729	83.729	79.901	79.901
Panel B: College Enrollment Rate (%)				
From High Achievement MS	1.751*	1.846*	3.301**	3.038**
	(0.967)	(1.018)	(1.542)	(1.444)
8th Grade ELA Score	1.314***		2.070***	
	(0.205)		(0.328)	
8th Grade Math Score	0.910***		1.416***	
	(0.231)		(0.374)	
Number of Programs Ranked		0.078		-0.320***
		(0.095)		(0.115)
N	7060	7060	6679	6679
R2	0.471	0.460	0.324	0.314
ȳ	72.197	72.197	67.204	67.204
Panel C: % High Performing Students				
From High Achievement MS	2.913*	3.567*	5.185**	5.232**
	(1.748)	(1.899)	(2.061)	(2.039)
8th Grade ELA Score	2.114***		3.023***	
	(0.351)		(0.409)	
8th Grade Math Score	1.258***		1.315**	
	(0.397)		(0.522)	
Number of Programs Ranked		0.500***		-0.050
		(0.119)		(0.122)
N	7062	7062	6751	6751
R2	0.513	0.510	0.415	0.400
ȳ	40.934	40.934	34.978	34.978

Notes: All columns show 2SLS estimates. Standard errors clustered at graduating middle school in parentheses, *p<0.1, **p<0.05, ***p<0.01. All regressions control for student ethnicity, gender, English Language Learner status, Free-Reduced Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. All columns also control for saturated dummy for all possible values of propensity score of being assigned to a high achievement MS, and local linear control for non-random tie-breakers.

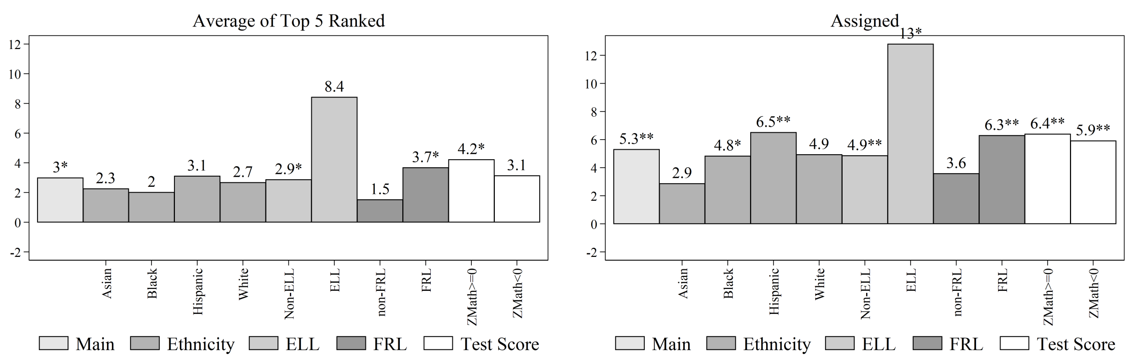
Table 2.18: Mediation Analysis

respectively. In a similar vein, attending a high achievement middle school has a larger effect for FRL students than their non-FRL peers and for Black, Hispanic, White students than their Asian peers. These results suggest that attending a high achievement middle school could *level the field* for different groups of students.

(a) College Enrollment Rate



(b) % High Performing Students

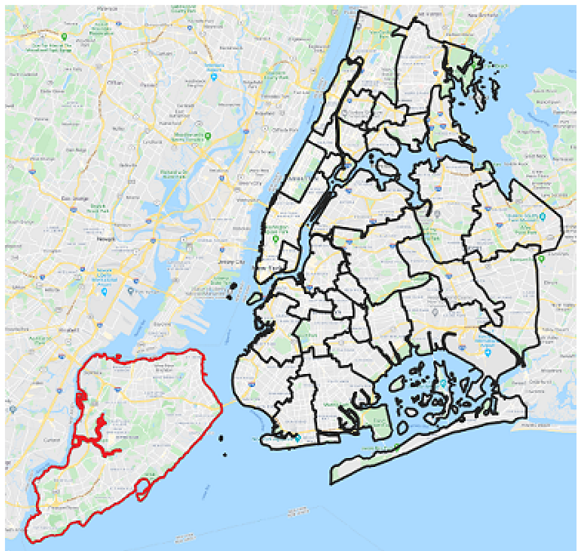


Notes: Standard errors are clustered at graduating middle school, *p<0.1, **p<0.05, ***p<0.01. We label a student 'high performing' if the standardized test score is above 66 percentile of the distribution. The estimates are derived by running 2SLS model (Equation 2.1) separately with students of the corresponding characteristics. Baseline student characteristics are controlled, excluding the demographic variable of interest. For instance, regression only with Asian students does not include the set of ethnicity dummy variables, but include ELL status, FRL status, and test score.

Figure 2.6: Effect of Middle School on High School Choice By Student Characteristics

Treatment Effect of Attending High Minority Middle Schools Next, we explore the treatment effects of attending a middle school with high fraction of Black or Hispanic students in Table 2.19. We do not find significant effects.

Additional Tables and Figures from Subsection 2.4.2 Figure 2.7 plots Staten Island whose students and schools we use for the model estimation.



Notes: The map shows 32 community school districts (CSD) in NYC. The red bordered is Staten Island, which is CSD 31 and well-separated with the rest of NYC.

Figure 2.7: Staten Island and NYC Community School Districts

Additional Tables and Figures from subsection 2.4.4

Alternative Assignments to Middle Schools Figure 2.8 reports the figures on decomposition of effects of middle schools on high school assignments, for two alternative counterfactual 'good' middle schools as described in Table 2.20.

When we exogenously change a student's middle school from a low achievement and high minority to a low achievement and low minority (Panel (a) of Figure 2.8),

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Average of Top 5			Matched		
Model	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Sample	All	All	NDR	All	All	NDR
Panel A: 4yr Graduation Rate (%)						
From High Minority MS	-1.306***	-0.549	-0.225	-1.480***	-0.034	-1.039
	(0.455)	(1.017)	(1.182)	(0.559)	(1.798)	(2.057)
N	46631	46631	3308	43927	43927	3103
R2	0.291	0.308	0.312	0.180	0.192	0.207
$\bar{y}$	83.438	83.438	79.593	79.097	79.097	74.068
Panel B: College Enrollment Rate (%)						
From High Minority MS	-1.686***	-0.314	0.248	-2.189***	-0.713	-0.794
	(0.553)	(1.267)	(1.459)	(0.661)	(2.052)	(2.383)
N	46630	46630	3307	43843	43843	3091
R2	0.363	0.378	0.358	0.237	0.250	0.260
$\bar{y}$	71.371	71.371	66.679	65.829	65.829	60.183
Panel C: % High Performing Students						
From High Minority MS	-4.024***	1.900	3.188	-3.875***	1.534	3.957*
	(0.850)	(1.745)	(2.084)	(0.800)	(2.046)	(2.240)
N	46723	46723	3317	44579	44579	3163
R2	0.441	0.455	0.370	0.376	0.387	0.333
$\bar{y}$	39.839	39.839	28.252	33.146	33.146	21.158
Panel D: % White						
From High Minority MS	-4.758***	-0.029	0.415	-4.152***	-0.930	-0.056
	(0.560)	(0.669)	(0.597)	(0.525)	(0.748)	(0.651)
N	46723	46723	3317	44579	44579	3163
R2	0.616	0.627	0.367	0.535	0.544	0.288
$\bar{y}$	18.518	18.518	7.045	14.957	14.957	4.242
Panel E: 1(STEM)						
From High Minority MS	0.034**	0.029	0.045	0.022	-0.036	-0.017
	(0.014)	(0.043)	(0.055)	(0.018)	(0.062)	(0.077)
N	46723	46723	3317	44582	44582	3163
R2	0.089	0.113	0.241	0.037	0.051	0.164
$\bar{y}$	0.325	0.325	0.376	0.315	0.315	0.361

Notes: Standard errors clustered at graduating middle school in parentheses, *p<0.1, **p<0.05, ***p<0.01. All regressions control for student ethnicity, gender, English Language Learner status, Free-Reduced Lunch eligibility, Special Education status, standardized test score in 5th grade, and residential borough in 5th grade. Column (2)-(3) and (5)-(6) control for saturated dummy for all possible values of propensity score of being assigned to a high minority middle school, and local linear control for non-random tie-breakers.

Table 2.19: Effects of Attending Highly Minority MS on HS Characteristics

	Middle School Types		Average Test Score
	High Achievement?	High Minority?	
Counterfactual 0		✓	602.01
Counterfactual 2			610.31
Counterfactual 3	✓		609.45

Notes: Counterfactual 0 is the same middle school as in Table 2.9. For Counterfactual 2 and 3, we choose the middle school that has the highest average score within each category among 'low achievement & low minority' and 'high achievement & low minority' respectively. See Table 2.9 for other details.

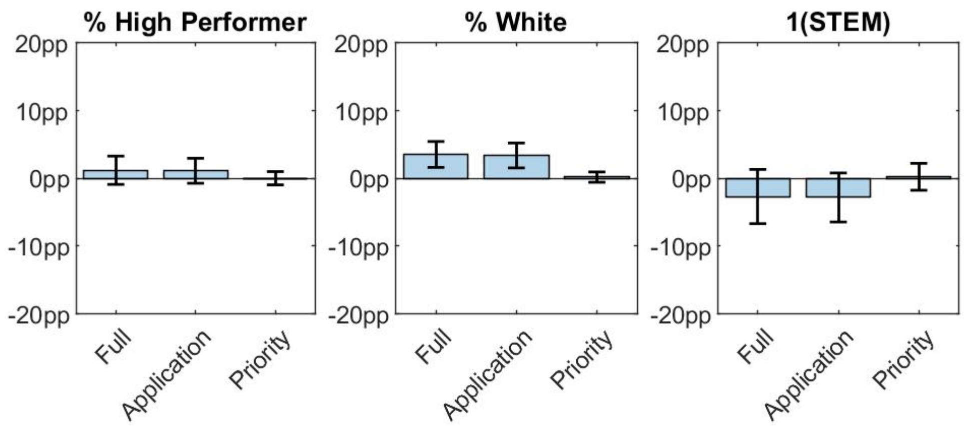
Table 2.20: Alternative Assignment to Middle Schools

we find that both the effects of the *application* channel and the *priority* channel are largely muted when compared to Figure 2.2, while the pattern that the relative magnitude of the *application* channel is larger is maintained. These findings align with what we find in the design-based approach that the effect of attending a middle school with a high fraction of Black or Hispanic is nearly null.

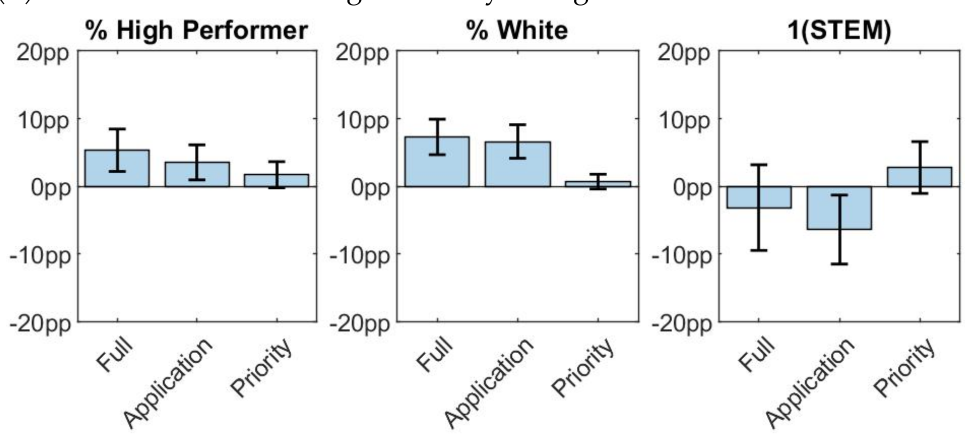
Next, when we exogenously change a student's middle school from a low achievement and high minority to a high achievement and low minority (Panel (b) of Figure 2.8), we find similar patterns but with a smaller magnitude.

Full Table

(a) Low Achievement & High Minority to Low Achievement & Low Minority



(b) Low Achievement & High Minority to High Achievement & Low Minority



Notes: See Figure 2.2 for the details on the figures.

Figure 2.8: Decomposition of Effects of Middle Schools on High School Assignments

		(1) Full	(2) Application	(3) Priority	(2)+(3) Application+Priority
Counterfactual 0 to 1	% High Performer	9.67	5.70	3.20	8.90
		(1.11)	(0.73)	(0.78)	
	% White	5.71	4.82	0.85	5.68
		(0.76)	(0.65)	(0.45)	
	1(STEM)	-0.54	-3.01	1.72	-1.29
		(2.11)	(1.55)	(1.44)	
Counterfactual 0 to 2	% High Performer	1.18	1.10	0.01	1.11
		(0.59)	(0.56)	(0.30)	
	% White	3.51	3.36	0.17	3.53
		(0.63)	(0.62)	(0.23)	
	1(STEM)	-2.70	-2.84	0.21	-2.63
		(1.25)	(1.20)	(0.57)	
Counterfactual 0 to 3	% High Performer	1.18	1.10	0.01	1.11
		(0.59)	(0.56)	(0.30)	
	% White	3.51	3.36	0.17	3.53
		(0.63)	(0.62)	(0.23)	
	1(STEM)	-2.70	-2.84	0.21	-2.63
		(1.25)	(1.20)	(0.57)	

Table 2.21: Decomposition of Effects of Middle Schools on High School Assignments

Additional Figures and Tables from subsection 2.4.5

Counterfactual Policy Predictions on Co-assigned Peers We report the results of counterfactual analyses on racial gap in the characteristics of the co-assigned peers.

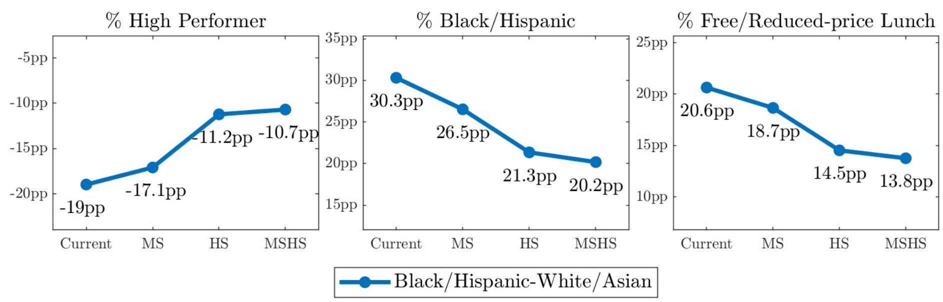


Figure 2.9: Racial Gap in Co-Assigned Peers in High School Choice

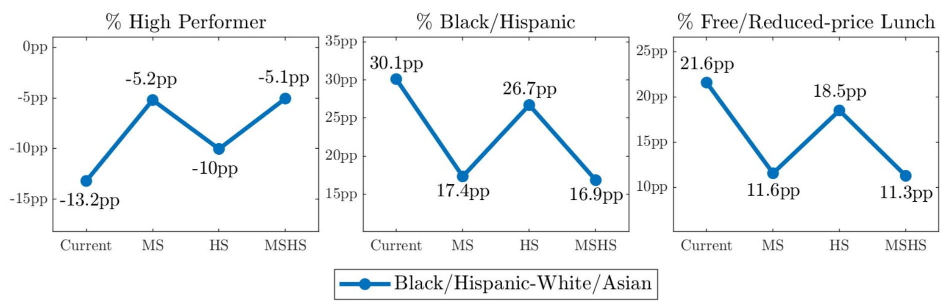


Figure 2.10: Racial Gap in Co-assigned Peers in Middle School Choice

3 MAKING COLLEGE AFFORDABLE? THE IMPACTS OF TUITION FREEZES AND CAPS

3.1 Introduction

In the face of concerns about college affordability, tuition freezes and caps are becoming an increasingly popular policy tool for state governments to regulate public colleges. They are a rare set of policies that often receive bipartisan support. Both parties frame freezes and caps as beneficial for state residents, who will be enabled to affordably obtain a college education.

A tuition freeze or cap occurs when a state government sets limits on the amount that public colleges are allowed to raise listed tuition (i.e. "sticker price") from year to year. Typically, a "freeze" occurs when colleges are banned from raising nominal tuition at all. However, states will frequently impose limits on the percent that colleges are allowed to increase tuition (e.g. 3 percent/year), rather than fully freezing tuition. From 1990 to 2013, seventeen states implemented a tuition freeze or cap at least once, affecting 2-3 percent of institutions and 7-8 percent of students each year (Deming and Walters, 2017). These tuition regulations typically only affect the in-state undergraduate tuition level.

Under an effectively enforced tuition regulation, colleges should not be able to increase listed tuition by a large amount. However, at the same time, they may search for other ways to compensate for their tuition losses. Such responses, in turn, can yield different results from what the state government intended by imposing a tuition regulation. Previous studies have found that colleges adjust various margins in response to different financial shocks (Dinerstein et al., 2014; Delaney and Kearney, 2015, 2016; Bound et al., 2020; Clelan and Kofoed, 2017; Deming and Walters, 2017; Webber, 2017), and which margin(s) colleges use under a tuition regulation is *a priori* ambiguous; they could decrease financial aid, hike up tuition once the regulation is lifted, or adjust other margins such as the composition of students by residency. Notably, depending on which margin universities adjust, tuition regulations could have different distributional implications.

Despite the prevalence of these policies and the *a priori* ambiguity in their effects, there has been little empirical evidence about the consequences of these tuition regulations. These effects are of direct interest to policy-makers considering these regulations, as well as to students and their families who may be subject to them.

Using a modified event study framework, we begin by estimating the effect of tuition freezes and caps on listed tuition to assess whether the regulations have any “bite". We find large heterogeneity in their effectiveness over time; tuition regulations have had large and statistically significant effects that have kept listed tuition from increasing in 2013 and earlier, but they have had no detectable effects from 2014 to 2019. We show that this is driven by a slowdown of tuition increases in recent years; in 2013 and earlier, institutions that were not under tuition regulations raised tuition by 6.3 percent, while the annual increase for these non-regulated institutions was 3.1 percent post 2014. This implies that colleges under tuition regulations are facing meaningful losses in tuition revenue in 2013 and earlier, but not in 2014 and later. Therefore, we expect to see colleges adjusting other margins such as institutional aid only in the earlier period, so we focus our analysis of outcomes other than listed tuition to the years before 2013.

Focusing on the earlier period, our primary finding is that although tuition caps and freezes reduce increases in “sticker price" tuition, they simultaneously induce universities to reduce increases in institutional financial aid, sometimes by a greater degree. This leads to an unintended consequence that when institutional aid is need-based, net benefit from a tuition regulation can be concentrated among richer students who do not receive institutional aid rather than needy students who do receive institutional aid. Dynamic changes in listed tuition and institutional aid over time have additional distributional impacts across cohorts, with some cohorts paying relatively higher tuition. Putting our results from the two periods together, our findings show that either these tuition regulations do not obtain their first-order goal of lowering listed tuition, or when they do, they simultaneously result in unintended distributional effects.

Specifically, we estimate that for four-year institutions, across all years a regulation is in place, the average yearly effect of a tuition regulation on listed tuition is -6.3 percentage points. To be precise, this means that listed tuition is 6.3 percentage points lower than it would have been in the absence of a regulation, and does not necessarily mean that tuition falls from year to year. All following effects should be interpreted in the same way. The corresponding impact on institutional aid is -11.3 percentage points. Two years after the end of the cap/freeze, listed tuition is 7.3 percentage points lower than it would have been if the regulation had never been in place while institutional financial aid is 19.5 percentage points lower. At two-year institutions, where the role of institutional aid is limited, colleges instead respond by rapidly increasing tuition once the cap/freeze has been lifted. During the regulation the impact on listed tuition is -9.3 percentage points; three years later it is only -4.8 percentage points and not statistically different from zero.

We probe for further heterogeneity in the four-year sector by estimating differential impacts by institution type. We find that institutions that are more dependent on tuition revenue lower financial aid more, and more quickly increase listed tuition after the regulation has been lifted. Similarly, we find that non-research universities adjust institutional aid more than research universities do.[1] These results imply that colleges with less monetary resources apart from tuition make larger adjustments to other margins in response to tuition regulations.

These responses from colleges imply that tuition caps and freezes have differential impacts on various groups of students. To give a sense of how this heterogeneity affects students moving through their education during and after a tuition regulation, we use our estimates to simulate the difference in net tuition paid from students' points of view. We consider students who vary in terms of 1) whether they receive institutional aid, 2) which type of institution they enroll in, and 3) when they first enroll with respect to the timing of the regulation. Our results imply that states that implement a uniform regulation on all colleges within the state may be creating inequalities in the way the regulation is felt by various students.

[1]A university is *More Dependent* if its fraction of total revenue from tuition and fees is greater than the median among public institutions. Research universities are defined as doctoral institutions with a Carnegie classification of high or very high research activity.

Depending on the type of student we consider, our estimates range from a student receiving a 5.9 percent discount to having to pay 3.8 percent more over four years of college due to the regulation.

While we focus on the effects on tuition and net tuition, we also extensively investigate effects on other outcomes such as room and board charges, instructional expenditure and the composition of students by residency and academic preparedness. We do not find any of them to be as important as adjustments in institutional aid, although we do find suggestive evidence that instruction-related expenditures per student are 3.3 percentage points lower under tuition regulations. These null results on other margins could be attributed to the fact that colleges are restricted in the changes they can make. For example, universities can not pool revenue from different sources when some part of the revenue is earmarked to pay for certain expenses by their budgeting practice or outside entities (Kelchen, 2016; Blagg et al., 2017). In 2010, 21% and 38% of total revenue of four-year and two-year institutions was restricted to be used for certain expenses.[2] Such restrictions can reduce incentives for increasing room and board prices, for instance, when universities can not shift the revenue to expenses sourced by tuition revenue.

Our paper fits into a literature investigating how colleges respond to financial shocks. Previous papers have studied implications of changes in state funding (Fethke, 2005; Webber, 2017; Bound et al., 2020) or federal funding (Singell Jr and Stone, 2007; Dinerstein et al., 2014). Among various outcomes, changes in listed tuition are often found to be the main channel through which universities adjust to financial shocks. For example, Webber (2017) finds that decreases in state funding are partially passed on to students through increases in tuition. He finds that on average between 1987 and 2014, students bear 25.7 percent of the financial burden from state funding changes. Similarly, Dinerstein et al. (2014) find that in response to the expansion of federal Pell grants during the Great Recession, public colleges raised tuition to fully capture the increase.

We study a different type of shock on colleges' revenue: tuition regulations. Tuition has been becoming a increasing share of universities' revenue due to steady

[2]Source: IPEDS

decreases in state funding in recent decades. Standing in contrast to other shocks, universities cannot adjust listed tuition to recoup the loss from the shock, by design of the regulation. We find institutional aid to be the most important margin that institutions adjust. While changes in tuition mostly yield distributional consequences from one cohort to another, changes in institutional aid can further result in an unequal distribution of benefits within each cohort.

Several papers have documented that universities often adjust institutional aid to capture additional revenue. Clelan and Kofoed (2017) find that institutional aid decreases during recessions. Turner (2017) shows that institutional aid is crowded out by federal Pell grants, with universities giving less institutional aid to students with higher Pell grants. In contrast to these papers that study a targeted policy (Pell grants) and a non-policy shock, we study a policy that is seemingly universal, at least among in-state undergraduate students. However, we show that even though tuition regulations are applied equally to all students paying in-state tuition, they can have different impacts across students because of institutions' responses of decreasing institutional aid.

Our paper also aligns with the small set of papers that focus on the tuition regulations specifically. Delaney and Kearney (2015, 2016) study impacts of the Illinois 2004 "Truth in Tuition" law, which requires flat tuition rates for 4 years for each cohort of students. They find that colleges increase tuition before cohorts enter in anticipation of not being able to increase it later. We use policy variation from all states in the US over the longer period (1990-2019) and find no anticipatory behavior but a statistically significant and economically meaningful response of changes in institutional aid in the first two decades. Relatedly, Deming and Walters (2017) exploit tuition freezes and caps in their analysis of whether increasing expenditures or lowering tuition is more effective in increasing enrollment and graduation at public colleges. They find a strong "first stage" effect of tuition caps/freezes on listed tuition; our results support this while adding the finding that the decrease in listed tuition is accompanied by decreases in institutional aid. This may be key to explaining the Deming and Walters (2017) finding that lower tuition (instrumented with tuition freezes) does not have a strong effect on total enrollment or graduation rates. We also add to this literature by examining heterogeneity in the type of regulation (i.e., cap or freeze and length of regulation) and university characteristics. Finally, we illustrate how this heterogeneity affects different types of students based on their timing of entry into college, the type of institution they attend, and whether they receive institutional financial aid.

The rest of the paper proceeds as follows: section 3.2 describes the institutional background and the data sets for our analyses, section 3.3 lays out a conceptual framework to frame empirical results, section 3.4 describes our empirical strategy and identification, section 3.5 presents results, section 3.6 illustrates the impact of a tuition regulation on a representative student by putting estimates together, and section 3.7 concludes.

3.2 Institutional Background and Data

The setting for our study is higher education institutions in the United States. Our primary analysis will be from 1990 to 2013, although we will show some specifications with more recent years (through 2019). We are interested in legislative tuition regulations and do not consider tuition freezes/caps initiated by colleges.[3] These tuition regulations almost exclusively affect only in-state undergraduate tuition; colleges are not regulated on how to set graduate tuition or out-of-state undergraduate tuition. Students fees are often regulated together with tuition, but financial aid is rarely regulated.[4]

These regulations are often put forth by politicians in an aim to make college more affordable for state residents. They are typically enacted as a part of the state higher education budget. This budget goes through multiple rounds of revisions. In addition to the general uncertainty of whether budget requests will be fully funded (which depends in part on tax revenues), there is uncertainty whether the tuition regulation will be enacted at the end of the budget process. There have

[3]For example, see Purdue University (2020).

[4]We found only one instance of tuition regulation packaged with institutional aid regulation (Rhode Island 2013-14 HB 7133, 2014-15 HB 5900).

been cases where either the upper house or the lower house of a state legislature proposes a bill for a tuition regulation but it does not pass the other house or the governor.[5] The duration a tuition regulation is often aligned with the duration of the budget bills because of this process; budget bills are sometimes done less frequently than annually such that a multi-year tuition regulation might be put into place. Tuition regulations could be extended to another fiscal year term when they are re-authorized along with the new budget bill, otherwise lifted. The uncertainly embedded in the budget approval process implies that it would be hard for an individual university to predict an upcoming tuition regulation.[6]

In this study, we will combine data sets from various sources. The main data is the Integrated Postsecondary Education Data System (IPEDS). IPEDS is a survey of colleges, universities and vocational institutions conducted annually by the U.S. Department of Education. All colleges that receive Title IV federal funding are required to report their data to IPEDS, so it is a universe of public colleges in the United States and a near universe of private colleges (aside from some for-profit institutions). IPEDS collects information on tuition and enrollment by student residency (i.e. in-state/out-of-state) status. IPEDS also collects detailed information on institutional finances and student financial aid, including revenues and expenditures by source.[7]

Our second data set is tuition regulations by state, detailing in which states and years tuition regulations were imposed. This data set, which we take from Deming and Walters (2017), distinguishes between tuition freezes and caps, and records the specific limits for tuition caps. In secondary analysis, we augment this data set

[5]E.g., Georgia 2016-17 HR 1326, Georgia 2018-19 SR 215, Tennessee 2014–16 HB 2179/SB 1683, Texas 2017-19 SB 19, Virginia 2018-19 HB 351).

[6]Our informal conversation with government relations officials at public universities indicate that tuition regulations are imposed with very little warning.

[7]We supplement our data with IPEDS finance data constructed and published by the Urban Institute (Blom et al., 2020). While the Delta Cost Project is well known to aggregate multiple institutions within some public university systems into a single administrative unit (Jaquette and Parra, 2016), the Urban Institute data leave that decision to the data user by reporting raw finance data and parent-child relationship among institutions (i.e., branches of a university system). In our analysis of state appropriations (presented in appendix Table 3.19), we do not aggregate parent-child observations.

by hand-collecting tuition regulations from 2014 to 2019 from state legislation. We collect this legislation through a combination of Lexis-Nexis searches of legislation and news articles, communication with state boards of education and legislatures, and verification using legislative records from state websites. We also double-check the data set from Deming and Walters (2017), making a few adjustments where we find discrepancies between their data and legislative records.

For our primary time period of focus, 1990-2013, 17 states imposed formal price regulations on public institutions at least once. For these 17 states between 1993 and 2013, 26.7 percent (109 out of 408) of state by year observations were under tuition regulations. In around half of these cases, institutions were under tuition freezes. The rest were tuition caps, with the exception of one case where institutions were mandated to cut tuition (Virginia, 2000). The caps ranged from three percent to 10 percent limits on increases in tuition. While some states imposed uniform price regulations on all public institutions, others differentiated by sector (see Table 3.8 and Table 3.9). Table 3.23 shows the full array of when and where freezes and caps were in place.[8]

Sometimes these regulations lasted for only one year, but they were often extended for multiple continuous years. When counting a regulation continued over multiple years as one regulation, 40% of regulations were lifted after one year (See Figure 3.1 for the whole distribution). Finally, Table 3.1 presents summary statistics of variables of interest by institution type (private/public, 4-year/2-year), with the first two columns showing statistics of institutions under tuition freezes or caps.

Our final two data sources consist of state level economic and political variables. First, we proxy for states' economic environments with unemployment rates from annual county level labor force data (U.S. Bureau of Labor Statistics, 2021). Second, we construct a variable indicating the majority party of each state's lower and upper legislative houses based on election data collected by Klarner (2018). This data covers each individual candidate who ran for state legislative office, with general election returns between 1990 and 2015, which we aggregate to the state by year

[8]Note that no regulations are in place in 1990 (our first year of data) so that we are starting with all "control" institutions.

level.

3.3 Conceptual Framework

Although we do not explicitly model a colleges' objective function, here we provide a general conceptual framework to provide context to our empirical results. Previous literature has shown that public universities do not necessarily act as profit-maximizing firms and that student characteristics (e.g. academic ability or socioeconomic status) can compose a main part of their objective function (Epple et al., 2006; Fu, 2014; Turner, 2017). Public universities maximize their objective function subject to a budget constraint. Our study focuses on responses to a change in the universities' ability to choose a key part of that budget constraint, namely, tuition. Diminishing state appropriations have made tuition revenue an increasingly important revenue source over the past 30 years.

In a given year, a college may optimally decide to increase listed tuition for several possible reasons. They may want to generate more revenue that can be used to increase quality (e.g. increase instructional expenditures). Alternatively, they may want to increase listed tuition while simultaneously increasing targeted financial aid so that they can enroll more students from the groups they care more about (e.g. high-ability students or low-income students). A tuition cap or freeze may force a college to deviate from its optimal tuition level. Still subject to a budget constraint, universities may seek to increase other revenue sources to recoup losses in tuition revenue. Part of these losses could be offset by more generous state funding. In our analysis, we see that being subject to a tuition regulation is associated with a 6 percentage point increase in state appropriations, which could be a result of negotiations between universities and state governments. But given that the tuition revenue is nearly one-third of total revenue on average, this might not be enough. Some universities could have other means such as donations, their endowments, or other university-run businesses, while other universities need to meet their budget constraint solely by decreasing expenditures.

Given this, we should expect to see adjustments along other margins such as changes in institutional financial aid or instructional expenditures. Which margin(s) will a university adjust? The answer could depend on many factors such as what other components construct the university's objective function, what other margins a state government regulates (e.g. number of out-of-state students), and the university's degree of market power in the higher education market. We neither explicitly model a college's objective function/the higher education market nor collect all information on other regulations. However, we interpret our results considering these factors, and furthermore our empirical results can shed light on universities' behavior.

Depending on which margin(s) universities adjust, how evenly impacts are distributed across students will vary. Some margins, such as changes in required student fees, could be expected to affect all students relatively evenly. Other margins may disproportionately affect certain groups of students. In the case of institutional aid, it is clear that students who receive institutional financial aid will be hurt more than students who pay "sticker price". In other cases such as instruction-related expenditures, the equity effect is more ambiguous and hinges on the relationship between universities' expenditures and it's heterogeneous effect on students.

3.4 Empirical Strategy

We use a modified event study framework to estimate the effects of tuition regulations on the dynamics of institutions' "sticker price" tuition and institutional financial aid. Together these two determine net price, which is more relevant than sticker price alone. Not all students receive institutional financial aid, so for some their change in net price will be equal to the change in sticker price tuition. However, for students who receive institutional aid, their change in net price will depend both on changes in sticker price tuition and on changes in institutional aid. Thus, if universities adjust financial aid in response to tuition regulations, they can have distributional impacts across students depending on whether they receive aid. We are also interested in the dynamics of how tuition and aid change during and after the regulations because these changes could differentially impact students depending on the timing of their college entry. For our benchmark specification, we estimate

$$y_{it} = \sum_{k=-3,k\neq-1}^{3} 1(\mathrm{TuitReg}_{t-k})_{it}\beta_k + \beta_4 \sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it} + \beta_{-4} \sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it} + \gamma_t + \phi_i + t\rho_{c(i)} + \beta_X X_{s(i)t} + u_{its} \quad (3.1)$$

where $1(\mathrm{TuitReg}_{t-k})_{it}$ is an indicator equal to 1 if institution i is under a tuition cap or freeze in year $t-k$ Observations more than 4 years before or after a tuition regulation are captured by $\sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it}$ and $\sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it}$, respectively.[9] γ_t is a calendar time fixed effect, ϕ_i is an institution fixed effect, $t\rho_c$ is a public/private-specific linear time trend, and X_{st} is a vector of time-varying state-year level controls. The control vector includes the state unemployment rate (along with its lead and lag) and the majority political party in each state's legislative lower and upper houses. Standard errors are clustered at the state level. We estimate Equation 3.1 separately for 2-year and 4-year institutions.

Our setup differs from a canonical event study set-up in two ways. First, an institution can be treated multiple times if a state has two (or more) separate tuition regulations during our time period, as opposed to being treated at most once as in a typical event study. To this end, we follow a strategy proposed in Sandler and Sandler (2014) which assigns a unique set of relative time indicators for each treatment. Under this strategy, the independent variables $1(\mathrm{TuitReg}_{t-k})_{it}, -3 \leqslant k \leqslant 3$ take values of at most 1 given that no two different tuition regulations could happen k years before a given year t when k is within three years of t. In contrast, $\sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it}$ takes values greater than one if, for instance, there was a

[9] In other words, we impose a constant coefficient for all periods 4 or more years before (after) the tuition regulation to deal with differential timing of tuition regulations; the only regulations for which we observe many pre- or post-periods are those with regulations at the tail ends of the data. Bailey and Goodman-Bacon (2015) use a similar strategy under a research design where an institution is treated at most once. Under such design, one can replace summations with the following dummy variables $1(\mathrm{TuitReg}_{t+k}, k \geqslant 4)_{it}$ and $1(\mathrm{TuitReg}_{t+k}, k \leqslant -4)_{it}$.

tuition regulation 4 years before a given year t and another regulation 7 years before t. The same holds for $\sum_{k=4}^{\infty} 1(\mathtt{TuitReg}_{t+k})_{it}$.[10]

Second, a tuition cap or freeze can last continuously for several years, with the length varying by state and enactment year. Figure 3.1 shows the distribution of the length of tuition regulations in our data. Whenever a tuition regulation lasts for more than one year, we impose a constant coefficient across all years in which the regulation was in place. Thus, the β_0 can be interpreted as the average yearly effect of the regulation across all years it was in place. This also implies that we can interpret the first lead as the year before the tuition regulation starts, and the first lag as the first year after the tuition regulation ends.[11]

While we expect β_k to be the weighted sum of treatment effects across tuition regulations at different timings, recent studies have shown that with heterogeneous treatment effects, the estimates from a two-way fixed effect (TWFE) regression might not be capturing this (De Chaisemartin and D'Haultfœuille, 2020b; Sun and Abraham, 2020). While this is a concern in our setting, newly proposed estimators that allow for staggered adoption either assume that treatment is an absorbing state (Callaway and Sant'Anna, 2020; Sun and Abraham, 2020) or abstract away from the dynamics (De Chaisemartin and D'Haultfœuille, 2020b). De Chaisemartin and D'Haultfœuille (2020a) extend their previous work to capture dynamic effects when the research design is non-staggered, albeit with some limitations; the proposed estimator captures the effect of switching into treatment k periods

[10]Therefore, β_4 is identified not only by the difference between treated and untreated units 4 and more periods after a tuition regulation but also by the linearity assumption on $\sum_{k=4}^{\infty} 1(\mathtt{TuitReg}_{t-k})_{it}$. In other words, the baseline specification assumes that the difference between a never-treated and a once-treated unit 4 or more periods after is same as the difference between the once-treated and a twice-treated unit after 4 or more periods. The same argument is applied to β_{-4}. To investigate if this linearity assumption matters, we run a variation of Equation 3.1 where we replace $\sum_{k=4}^{\infty} 1(\mathtt{TuitReg}_{t-k})_{it}$ with a set of dummy variables $1(\sum_{k=4}^{\infty} 1(\mathtt{TuitReg}_{t-k}) = N)$. Our coefficients of interest, the β_ks, $k = -3, -2, 0, 1, 2, 3$, are very robust with the modification.

[11]Note that if all tuition regulations had the same length, we could follow a more conventional event study specification by separately estimating a coefficient for each year of the regulation. This is infeasible due to large variation in the length of tuition regulations. However, because we are interested in dynamic effects of the regulation over the period in which it is in place, we estimate additional specifications that explicitly incorporate coefficients to capture variation in the length of regulations (see equations Equation 3.4 and Equation 3.5 below).

ago, averaging different trajectories of treatment histories afterwards. Given these limitations, we use the two-way fixed effect design in Equation 3.1 as our baseline specification. However, we estimate our treatment effect using the estimator from De Chaisemartin and D'Haultfœuille (2020a), but focus only on the instantaneous effects to avoid averaging different treatment histories. The results are presented in appendix Table 3.16, and are closely aligned with the estimates from our baseline specification.

We include a public/private-specific linear time trend rather than a state-specific trend in our main specification for two reasons. First, the inclusion of the public/private-specific trend helps us meet the parallel trends assumption while we see a positive pre-trend in the state-specific trend specification. Moreover, there could be spillover effects on private colleges located in the same state; private colleges could set their tuition or aid taking those of their competitors into account. For instance, Epple et al. (2006) study how colleges set listed prices and institutional aid in an equilibrium setting. We will show some evidence of spillover effects in subsection 3.5.7. We also present sensitivity analyses where we instead use sector-year fixed effects or a state-specific linear time trend in appendix Table 3.17.

The coefficients of interest are the β_ks with $k = 0, 1, 2, 3$.[12] When the outcome variable is listed tuition, β_0 measures how effectively the tuition regulation was enforced whereas the β_ks with $k = 1, 2, 3$ capture how colleges adjust tuition after the regulation has ended. With other outcomes (e.g. institutional aid), the β_ks with $k = 0, 1, 2, 3$ show how colleges adjust other unregulated margins during and after tuition controls.

With our normalization which omits $1(\mathtt{TuitReg}_{t-1})$ in Equation 3.1, β_k captures the additional difference in y_{it} between treated and untreated units k periods after[13] the tuition cap or freeze is imposed, beyond the difference in the -1 period (which

[12]We do not focus on β_{4+} since its interpretation is unclear due to the aggregation of periods and differing amounts of observations at the tail ends of the time period studied.

[13]In the case where $k < 0$, this can be interpreted as $-k$ periods before the treatment. For example, $k = -2$ implies it is two years before the treatment.

has been normalized to zero). In equation form,

$$\beta_k = E(y_{it-k}|R=1,\tilde{X}) - E(y_{it-k}|R=0,\tilde{X}) \\ - \left(E(y_{it-k-1}|R=1,\tilde{X}) - E(y_{it-k-1}|R=0,\tilde{X})\right) \quad (3.2)$$

where $R = 1$ is a university with a tuition regulation k periods before $(1(TuitReg_{t-k})_{it} = 1)$ and $R = 0$ is a university without a tuition regulation. In addition, $\tilde{X}$ represents the collection of γ_t, ϕ_i, $t\rho_c$ and X_{st} from Equation 3.1. We can interpret β_k, $k \geqslant 0$ as a causal effect of a tuition cap or freeze only when the parallel trends assumption holds, i.e., the mean change in the unobserved part of treated observations over time is equal to that of untreated observations after conditioning on $\tilde{X}$.

To bolster the case for a causal interpretation, we do three things. First, we investigate coefficients β_k, $k < 0$, in the years prior to the tuition regulation. It's possible that the state government could use the regulation as a punishment for colleges that have been increasing tuition rapidly. On the contrary, they could take advantage of colleges that are already slowing down tuition increases by advertising the tuition regulation to voters without having any meaningful impact on tuition setting. However, in these cases, we should see this behavior in the years leading up to the tuition regulation. We do not see evidence of this, as the values of β_k are not statistically different from zero when $k < 0$. If anything, we see a small pre-trend upward in the years leading up to the regulation for both listed tuition and financial aid, so adjusting for this would strengthen our main results.

Second, we control for several key variables in Equation 3.1. Institution fixed effects capture any non-time-varying differences between treated and untreated units. Our public/private-specific linear time trend captures a linear approximation of time-varying differences between private and public schools. The calendar time fixed effect captures the national-level time trend. Our inclusion of state-level unemployment rates, and their leads and lags assuage concerns about the Great Recession or other state-varying macroeconomic trends affecting results.[14]

[14]We use labor force data by county from Local Area Unemployment Statistics (LAUS) announced annually by Bureau of Labor Statistics (BLS). We control for the average unemployment rate by

Finally, we include indicators for the majority political party to capture state-varying differences in political factors that may affect both tuition prices and the probability of a state imposing a freeze/cap.

Third, we implement robustness checks with different comparison groups. First, we have a specification that only includes institutions that have been under a tuition cap or freeze at least once during the time frame studied. In this analysis, we leverage only variation in the timing of cap/freeze, exploiting the fact that different states imposed tuition regulations at different times (Bailey and Goodman-Bacon, 2015). Second, we implement a matching procedure where we match treated institutions to untreated institutions with similar tuition levels and trends in the years prior to the regulation.

Conceptually, we are thinking of the results we see as colleges' response to a tuition cap or freeze being imposed on them. However, there are cases where we want to be cautious with this interpretation. First, we might be picking up other policies imposed on colleges that happen at the same time as the tuition regulation. Specifically, states imposing tuition caps/freezes often simultaneously give more generous funding to colleges as compensation. Our analysis show that institutions have 6 percentage points higher state appropriations during a tuition regulation (this effect is not statistically significant for four-year institutions but significant at a 5 percent level for two-year institutions. For more detail, see appendix Figure 3.9). In this case, our coefficient would capture the combined effect of the cap/freeze and the state funding. Thus we implement a sensitivity check where we control for state funding, and our findings of the effect of tuition regulations on tuition and aid are robust (see appendix Table 3.19).[15]

Moreover, state governments may be aware of changes in the unobservable u_{it}

state, aggregated from counties within each state weighted by the size of labor force population.

[15]We do not control for state funding in our main specification because state appropriations could be determined as an outcome of the negotiation between colleges and the state after a tuition cap/freeze is imposed. In this case, colleges with different unobservable characteristics such as their bargaining power could select into different levels of increases in state funding. (This is a "bad control" discussed in Angrist and Pischke (2008) in detail. Webber (2017) also uses a sparse set of time-varying controls for the same reason in a similar context to ours.). However, results from our robustness check show that this might not be a concerning issue in our context.

and use it to make a decision of whether to impose a tuition regulation. Previous work has shown that state governments adjust appropriations based on temporary financial shocks to colleges (Dinerstein et al., 2014; Fu et al., 2019). It is also possible that the state government and colleges could be jointly deciding whether to have a tuition regulation. In this case, our estimates would simply show what happens during and after a tuition regulation. Notably, our interpretation of effects on students (and how effects vary with student characteristics) remain the same.

In addition to the benchmark specification in Equation 3.1, we run two other specifications. First, we explore heterogeneity in whether schools experience a freeze or a cap (and in the size of the cap). Specifically, we estimate

$$y_{it} = \sum_{k=-3,k\neq-1}^{3} 1(TuitReg_{t-k})_{it}\beta_k + \sum_{k=0}^{3}(TuitCap_{t-k})_{it}\alpha_k$$
$$+ \beta_4 \sum_{k=4}^{\infty} 1(TuitReg_{t-k})_{it} + \beta_{-4} \sum_{k=4}^{\infty} 1(TuitReg_{t+k})_{it} + \phi_i + \gamma_t + t\rho_c + X_{st} + u_{its} \tag{3.3}$$

which is the same as our benchmark specification except in the second term. $(TuitCap_{t-k})$ represents the size of the cap and is coded from 0 to 1; for a 3 percent cap, $(TuitCap_{t-k}) = 0.03$. When tuition is frozen, $(TuitCap_{t-k})$ takes a value of 0. With this specification, β_k represents the effect of tuition being completely frozen. The effect of tuition cap is $\beta_k + \alpha_k \times (TuitCap_{t-k})$.[16]

We also run regression models that consider the variation in the length of tuition regulations.

[16]We do not include a tuition cap coefficient for the endpoint coefficients since their interpretations are unclear due to differing amounts of observations at the tail ends of the time period studied.

$$y_{it} = \sum_{k=-3,k\neq-1}^{3} 1(\mathrm{TuitReg}_{t-k})_{it}\beta_k + 1(\mathrm{FirstYrofTuitReg}_t)_{it}\alpha_F + 1(\mathrm{LastYrofTuitReg}_t)_{it}\alpha_L$$
$$+ \beta_4 \sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it} + \beta_{-4} \sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it} + \phi_i + \gamma_t + t\rho_c + X_{st} + u_{its} \tag{3.4}$$

$$y_{it} = \sum_{k=-3,k\neq-1}^{3} 1(\mathrm{TuitReg}_{t-k})_{it}\beta_k + (T_{it} - 1)\alpha_A$$
$$+ \beta_4 \sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it} + \beta_{-4} \sum_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it} + \phi_i + \gamma_t + t\rho_c + X_{st} + u_{its} \tag{3.5}$$

Equation 3.4 additionally includes indicators for the first and last year of the cap/freeze. $1(\mathrm{FirstYrofTuitReg}_t)_{it} = 1$ if the institution is under the first year of tuition cap/freeze. $1(\mathrm{LastYrofTuitReg}_t)_{it}$ is defined similarly. In this specification, β_0 gives the average effect for all years other than the first and last year in which the regulation is in place. The value of the outcome variable in the first/last year of tuition regulation is equal to $\beta_0 + \alpha_F$, $\beta_0 + \alpha_L$, respectively.[17]

Equation 3.5 allows each additional year of a tuition regulation to have a linear effect on tuition and fees. T_{it} represents the number of consecutive tuition regulations up to year t. Thus, β_0 represents the effect of having a tuition regulation in place for exactly one year. The effect of having a tuition regulation for 5 years continuously is given by $\beta_0 + \alpha_A \times (5 - 1)$.

[17] If a tuition cap/freeze lasts only one year, both the first and last year dummy variables are switched on. If it lasts for two years, the first year is switched on for the first year and the last year for the second year.

3.5 Results

3.5.1 Effects on Listed Tuition during Regulations

We begin by investigating the first-order effects of tuition freezes and caps on listed tuition while they are in place. Since this is the outcome being directly targeted by the policy, we consider this outcome to be a measure of whether the tuition regulation has “bite". Although our main outcomes of interest will be how colleges adjust unregulated margins during and after tuition regulations, we would only expect to see these adjustments when the regulation has some bite. To test this, we estimate the full specification in Equation 3.3 on the log of in-state undergraduate listed tuition and fees but present only our estimates of β_0 and α_0 to focus on the contemporaneous effect while the regulation is in place.[18] Table 3.2 shows the results and reveals that tuition regulations only had bite in the earlier time period. The first row shows that across all institutions, having a tuition freeze has an approximately -11.2 percentage point impact[19] on tuition in the years 2013 and earlier, but no impact in the years 2014 and later. This pattern persists when separately estimating effects for four-year and two-year institutions. The second row shows the effects of an institution having a tuition cap rather than a freeze. In the earlier period for all institutions, each additional percentage point in the cap reduces the effect by around 1 percentage point. For example, the effect of a 5 percent tuition cap is -11.2 + 0.05(100.1) = -6.2 percentage points.

Table 3.3 illuminates one of the driving forces by comparing the average annual increase in tuition between treated and untreated institutions over the two time periods. In 2013 and earlier, institutions under tuition regulations raised tuition by 2.4 percent each year on average while institutions not under tuition regulations

[18]We do not yet want to consider impacts on tuition in the years after the regulation since these could be capturing the response of colleges to tuition regulations once they regain control of tuition setting.

[19]This interpretation comes from the following calculation. Note that we use log of tuition. $\beta_0 = -0.112$ means $E(\log\frac{P_t}{P_{t-1}}|1(TuitReg_t)_{it} = 1) - E(\log\frac{P_t}{P_{t-1}}|1(TuitReg_t)_{it} = 0) = 0.112$. Using the approximation that $\log(1 + x) \approx x$ when x is small, we have $E(\frac{\Delta P_t}{P_{t-1}}|1(TuitReg_t)_{it} = 1) - E(\frac{\Delta P_t}{P_{t-1}}|1(TuitReg_t)_{it} = 0) = -0.112$, where $\Delta P_t = P_t - P_{t-1}$.

raised tuition by 6.3 percent. Since 2014, treated institutions have behaved similarly as before, raising tuition by 2.5 percent each year. However, institutions that were not regulated only raised tuition by 3.1 percent, less than one percentage point above the treated group. Because institutions that were not forced to keep tuition levels down were not raising tuition much, the tuition freezes and (even more so) caps had essentially no bite.[20][21]

Our primary interest is on the downstream effects of these regulations. That is, how institutions respond by changing margins they still control, such as institutional aid. When the tuition regulation is not so effective in lowering listed tuition, colleges do not have adjust to make up for the loss from the regulation, so we would not expect to find effects on other margins. Therefore, in the rest of out analysis we focus on the time period of 1990-2013 to understand how universities respond to tuition regulations when they have some bite.

3.5.2 Dynamics of Listed Tuition and Institutional Aid

Figure 3.2 shows results of having a tuition regulation (either cap or freeze) by estimating Equation 3.1 for two outcomes: log of in-state undergraduate tuition and fees[22], and log of institutional financial aid for first-time undergraduates students. The solid lines represent coefficient estimates and the dotted lines represent 95 percent confidence intervals. Focusing first on four-year colleges in panel (a), we see that neither in-state tuition nor institutional aid statistically differs from zero in most years prior to the tuition regulation. This evidence supports our parallel trends

[20]These results are not sensitive to the specific year we choose to cut the data within the years between 2009 and 2014. We decide to use 2013 as a cutoff for our main results since this is where we switch from using Deming and Walters (2017) data to our own hand-collected data, and although we tried to follow their methods there may be some differences in collection procedures.

[21]Although it goes beyond the scope of this study to understand the causes behind the slowdown of tuition increase in recent years, one conjecture is that there has been increasing attention on the price of higher education, which often results in negative media coverage or political discussion on tuition "hikes".

[22]Results using tuition levels rather than the log of tuition are similar and can be found in Table 3.12. We use the sum of tuition and fees because this variable is available for the entire time period we study whereas tuition alone is not available until 2000.

assumption, which requires that there are no effects of having a tuition regulation in the future, because at this point, neither group has experienced treatment yet. If anything, both tuition and aid are slightly increasing in the years prior to tuition regulation so adjusting for this trend would make decreases in the years following tuition regulation larger.

Next, we are interested in the coefficient at period 0, which gives the effect of a tuition regulation on tuition and fees while the regulation is in place. As expected, we see a statistically significant negative effect (-6.3 percentage points).[23] One year after the regulation has been lifted, we still see a negative effect on tuition of 8.5 percentage points, which is slightly larger than the effect during the cap/freeze. This is due to the fact that the coefficient at period 0 captures the average effect over multiple years of tuition regulations.

To further understand the dynamics of tuition regulations that last for more than one year, Figure 3.4 illustrates the results from Equation 3.4. In this plot, "First Year" gives the effect of the tuition regulation on in-state tuition and fees in the first year that the regulation is in place, "Last Year" gives this same effect in the final year the regulation is in place, and "Middle Years" give the average effect for all years other than the first and last year in which the regulation is in place. The figure shows that as tuition regulations last longer, their cumulative impact on the amount that tuition and fees deviates from its trend becomes larger, with a -2.2 percentage point estimate in the first year and a -11.6 percentage point estimate in the final year for four-year colleges. The easiest way to think about this is in the context of a three-year regulation, where tuition steady falls further from the trend in each of the years. If, instead, it was a four-year regulation, the "Middle Years" would represent the average of the second and third year, and so on with longer regulations. These differences are statistically significant: we can reject a null of a constant treatment effect between the first year, middle years, and last year with a p-value less than 0.001. In a similar vein, columns 2 and 4 in appendix Table 3.11

[23]This does not match our estimate from Table 3.2 because here we use the specification without heterogeneity between freezes and caps for ease of interpretation in the figures. Results incorporating this heterogeneity are discussed below and can be found in Table 3.4 and Table 3.10.

present results from Equation 3.5. The effect of having a tuition regulation in place for exactly one year is -2.3 percentage points. Having another consecutive year of regulation lowers tuition by an additional 9.9 percentage points. These results support the conclusion that the cumulative effect of tuition regulations increases as the regulation lasts longer.

Continuing to focus on the years after the regulation is lifted, both Figure 3.2 and Figure 3.4 show that tuition remains lower than it would have been in the absence of the regulation for three years after the end of the cap/freeze, with some evidence of small increases as institutions "catch up" to where they would have been without the regulation.[24] The absence of a faster catch-up may be related to state variation in the degree of autonomy that institutions have to set tuition rates, as noted by Webber (2017). All of the coefficient and standard error estimates for Figure 3.2a and Figure 3.4a can be found in Table 3.4 and Table 3.11, respectively.[25]

Panel (b) of Figure 3.2 and Figure 3.4 show these patterns for two-year colleges. The patterns in both figures are similar to those of four-year institutions, although the magnitudes are bigger: the effect on tuition is -8.2 percentage points on average during the regulation and -18.7 percentage points in the last year of the regulation.[26] Despite the larger negative effects of the tuition regulation on tuition during the cap/freeze, we see a much stronger "catch up" effect for two-year institutions.[27] By the third year after the freeze/cap ends, there is no statistically significant difference between actual tuition and counterfactual tuition in a world where the college did not experience any cap or freeze. We suspect that two year colleges exhibit a stronger "catch-up" effect than four-year colleges because two-year colleges have less room to adjust along the institutional aid margin, given that initial levels of institutional aid at two-year colleges are very low, as presented in Table 3.1. All of

[24]A joint test of equality between the coefficients 1, 2, and 3 years after the regulation can be rejected with a p-value of less than 0.001.

[25]In Table 3.11, the effect of the first year of the tuition regulation is $1(\mathtt{TuitReg_t}) + 1(\mathtt{FirstYRofTuitReg_t})$, while the effect of the last year is $1(\mathtt{TuitReg_t}) + 1(\mathtt{LastYRofTuitReg_t})$.

[26]Similar to 4-year schools, we can also reject a test of constant treatment effects during the regulation with a p-value of less than 0.001.

[27]We can reject a constant treatment effect among the last year of the regulation and the first, second, and third year after the regulation with a p-values less than 0.001.

the coefficient and standard error estimates for Figure 3.2b and Figure 3.4b can be found in appendix Table 3.10 and Table 3.11, respectively.

The line with triangle marks in Figure 3.2 shows the effect on institutional financial aid during and after the tuition regulation. Institutional aid includes all grants given by the university to students, and does not include loans or any financial aid that the student receives from the government or any other source outside the institution. Colleges decrease institutional aid by a greater proportion than tuition, which suggests that they use institutional financial aid as a way to recoup some of the tuition losses from the tuition regulation. The pattern of institutional aid in the years after the regulation follows a similar path to that of tuition, although always of lower magnitude. The difference in the effect on institutional aid and the effect on tuition is statistically significant at the 5 percent level in every year following the regulation and at the 10 percent level during the regulation.[28] As a result of college's response of decreasing institutional aid, we indeed find that the average net tuition does not decrease significantly neither during nor following the tuition regulation (Figure 3.3). Because institutional aid is unlikely to be a large factor at two-year colleges, we do not include estimates for institutional aid in panel (b).[29]

There are two other possible explanations worth mentioning for the negative effect on institutional aid. First, students are spending relatively less on tuition, so they should need a smaller amount of aid to cover their costs. Relatedly, it could be that institutional aid decreases mechanically following the decrease in tuition if the amount of the aid is tied with the amount of tuition (e.g. aid is X percent of tuition). However, we see that the magnitude of the effect on institutional aid is not only bigger during the tuition cap/freeze, it falls further after the regulation is lifted.

Second, tuition regulation could change the composition of students that institutions enroll. This could make the new student body different in terms of income or

[28]Results come from a GMM setup with conditions derived from two event study regressions, one with tuition as the outcome variable and the other with institutional aid. The p-values on the difference of the two effects 1, 2, and 3 years after the regulation are 0.028, 0.006, and 0.020, respectively. The p-value is 0.055 for the difference of the effects during the regulation.

[29]However, estimates can be found in appendix Table 3.10.

academic preparedness, which could explain a change the amount of aid. However, Figure 3.8 shows that federal Pell grants and state grants to students were not affected by tuition caps/freezes. Given that Pell grants are need-based, this suggests regulations didn't lead to a big change in the student composition by income. Like institutional aid, state aid is awarded by both need and merit. We do not see a clear effect of tuition regulations on state aid either.[30] Further, appendix Table 3.22 shows there is no effect of tuition regulations on first-time students' SAT scores, giving more direct evidence that colleges' student composition by academic preparedness did not change. These results support our interpretation that the negative effect on institutional aid is at least in part an effort by institutions to make up for lost tuition revenues.

Table 3.5 illustrates the dynamics of tuition revenue in response to tuition regulations. During a regulation, both gross and net tuition are lower than they would have been in the absence of the regulation. However, this negative effect is over two million dollars larger for gross tuition revenue (-4.7 million dollars, statistically significant at 10%) than for net tuition revenue (-2.7 million dollars, not statistically significant). This adds to our evidence that colleges decrease institutional aid to make up for tuition losses. After the regulation is lifted, the effects on both gross and net tuition revenue are no longer statistically significant (although still sizeable).[31]

To give a sense of the impacts of tuition regulations in dollar terms, we present results with the outcome variable as levels of tuition and fees (as opposed to logs) in Table 3.12. Column (1) shows that a tuition regulation has a -268 dollar effect on in-state tuition and fees each year during the regulation. Column (3) shows that colleges are almost completely compensating for this loss with institutional aid: the effect on aid is -212 dollars each year. Institutional aid continues to lag

[30]These results also show that the decrease in institutional aid was not offset by any increases of state aid or Pell grants.

[31]Given that revenue is tuition times the number of students, we check if there is an effect of tuition regulations on the total number of enrolled students but find no evidence of this. The coefficient of $1(\mathtt{TuitReg}_t)_i t$ is -23 with robust standard error 165.55 when we regress a measure of full time equivalent students on dummies of tuition regulations and control variables.

behind where it would have been in the absence of a cap/freeze in the years after the cap/freeze has ended, even more than tuition in some years.

In addition to representing the information conveyed in the figures described above, columns 2 and 4 of Table 3.4 and Table 3.10 present estimates from Equation 3.3 where we differentiate tuition caps and freezes. Focusing first on four-year colleges in Table 3.4, we see that the effect of a 5 percent tuition cap is -9.4 + 0.05(96.7) = -4.6 percentage points. When tuition is frozen, $(\mathsf{TuitCap}_{t-k})$ takes a value of 0, so the coefficient of -0.094 indicates that the effect of tuition being completely frozen on in-state tuition and fees is -9.4 percentage points for each year that is it frozen. This specification shows the intuitive result that institutions under caps experience smaller negative effects on tuition than institutions under freezes during and after the regulation. Three years after the end of the regulation, the tuition at colleges that had a freeze are still 9.4 percentage points behind where they would have been without the freeze. Meanwhile, those with a 5 percent cap are only 5 percentage points behind. The patterns for institutional aid at four-year colleges, as well as tuition at two-year colleges shown in appendix Table 3.10, are similar.

3.5.3 Heterogeneity

Next, we investigate heterogeneity in four-year colleges' responses to tuition freezes. First, we look into whether colleges' dependency on tuition affects how they respond to tuition regulations. Following a strategy of measuring state appropriations dependency from Deming and Walters (2017), we categorize institutions into more or less dependent on tuition based on the fraction of their total revenue that is sourced from tuition and fees in the initial year of our data, i.e. 1991. If this fraction is greater than the median fraction for all public institutions, the institution is classified as *More Dependent* whereas institutions with a fraction less than the median are classified as *Less Dependent*.

Figure 3.5 shows the results. Focusing first on in-state tuition (grey lines with circle markers), we see that institutions that are more dependent on tuition seem to increase tuition faster in the years following the end of the regulation, presumably because they do not have as many other sources of revenue to pull from when they take a loss from the tuition regulation. Similarly, institutions that are more dependent on tuition decrease their institutional aid more during and following the tuition regulation. These results support our interpretation of the decrease in institutional aid in our main results as being due to colleges adjusting to make up for tuition revenue losses.

Next, we break down institutions into three broad categories from the Carnegie classification system, using a modification of the classification from Bound et al. (2019). *Research* universities are doctoral-granting universities with high or very high research activity. The *Non-Research* group includes masters-granting universities and doctoral-granting universities with low research activity. All other 4+ year degree granting institutions fall into the *Other* category.

Figure 3.6 reveals that although the coefficients on tuition during the time of the regulation were of a similar size, there are differences in the tuition-setting behavior of colleges in the years following the cap or freeze. The *Non-Research* and *Other* groups seem to "catch up" a little more quickly while the *Research* universities' tuition remains well below where it would have been in the absence of the regulation. This may be because *Research* universities have more resources and do not need to raise tuition as rapidly to make up for the losses incurred by the regulation.

More strikingly, there is a discrepancy among the way these groups of colleges adjust their institutional aid. *Research* universities seem to reduce institutional aid in proportion to the reduction in tuition during the regulation and in the first year following, but then increase it slightly in the next two years. *Non-Research* universities do not adjust much during the regulation but reduce institutional aid in a proportion greater than tuition in the years following the cap or freeze. Finally, *Other* institutions have a sharp decline in institutional aid offered during the regulation that remains below the reductions in tuition for several years after the end of the regulation.

A possible mechanism for the heterogeneous responses by Carnegie classification come from the fact that the classification is a proxy of the university's available resources as well as stature and selectivity. *Non-Research* and *Other* universities are more dependent on tuition revenue than *Research* universities. While 32% of total revenue is sourced by tuition in *Research* universities on average, 59% and 54% of revenue is for *Non-research* and *Other* universities, respectively. Moreover, selective universities could leverage the higher demand from students to find ways to compensate their losses from tuition regulations. For example, Bound et al. (2019, 2020) find that facing a steady decrease of state funding, *Research* universities admit more out-of-state and foreign students who pay higher tuition than in-state students. Of course, our heterogeneity results by Carnegie classification should be taken with caution due to the large standard errors associated with the coefficients on institutional aid.

3.5.4 Robustness

In this section, we perform five analyses to ensure the robustness of out results. First, we implement a matching procedure to ensure that treated and comparison units are balanced on their tuition levels and trends before the regulation is put into place. Matching results can be found in the first two columns of Table 3.15. We implement 1-1 matching of institutions by year based on the Mahalanobis distance of the level of in-state undergraduate tuition and the annual rate of increase in in-state undergraduate tuition for the years one, two, and three years before the tuition cap/freeze.[32] The main conclusions from our baseline analysis remain.

Second, we include a specification that only includes institutions in states that were treated at some point during the time period we study. This is motivated by a potential concern that there may be some unobserved differences between the time trends of states that are subject to tuition regulations and states that never experience a tuition regulation. This version leverages only variation in the timing of the tuition regulations, rather than both the timing and existence of tuition regulations. Columns 3 and 4 of Table 3.15 restricts the sample to “ever treated" institutions. Although estimates are nosier than our main results, the signs and magnitudes of estimates are very similar.

[32]We use the user-written Stata command *kmatch* (Jann, 2017).

Third, we limit the sample to only observations where we observe both tuition and aid, which changes the sample dramatically since institutional aid data does not become available until 2001. This helps us ensure that the relative magnitude of "sticker price" and institutional aid is not driven by differences in estimating samples. The final column of Table 3.15 shows results for in-state tuition when only including observations that are in our estimating sample for institutional aid. Our results are robust and if anything indicate a greater gap between the change in in-state tuition and institutional aid.

Fourth, motivated by the recent literature showing pitfalls of TWFE estimators (De Chaisemartin and D'Haultfœuille, 2020b; Sun and Abraham, 2020), we estimate the effect of tuition controls on the listed tuition and aid using the estimator proposed in De Chaisemartin and D'Haultfœuille (2020a). This estimator captures the effect of the first time switching into treatment k periods ago, averaging the effect of different trajectories between $t - k - 1$ and the observation year t. To ease interpretation, we present the effect of an institution's first tuition regulation on its listed tuition and institutional aid during the first year that it is under the regulation in Table 3.16.[33] The estimator uses not-yet treated observations up to the year t as the comparison group. The estimates support the main story from the our baseline TWFE specification; 1) the magnitude of the effect on aid is greater than that on listed tuition (Tuition: -0.035 vs. Aid: -0.11), 2) universities that are *More Dependent* on tuition revenue adjust aid more (*More Dependent*: -0.206 vs. *Less Dependent*: -0.019), and 3) there is large heterogeneity in treatment effects by Carnegie classification.

Finally, there may be some concern that our estimates are picking up not only the effects of tuition regulations, but the combined effect of tuition regulations and changes in state and local funding. To address this, we investigate the relationship between state and local funding and tuition caps/freezes. Although we find that during a tuition regulation, institutions receive 6 percentage points more in state

[33]Note that this captures a different effect than our baseline specification which presents the average yearly effect across all years in which the tuition regulation was in place. It is more similar to our estimated effect of the first year of a regulation from Equation 3.4.

appropriations (not statistically significant for four-year colleges), if we control for state and local funding in our main specification, the coefficients of interest do not change. Appendix Figure 3.9 shows the estimated effect of a tuition regulation on state funding. Appendix Table 3.19 shows estimates of the effects of tuition regulations on tuition and institutional aid after controlling for state and local funding. Columns 1 and 2 give effects for four-year institutions, while column 3 shows results for two-year institutions.

3.5.5 Effects on Expenditure

In addition to adjusting revenue (i.e. net tuition) in response to financial shocks, colleges may also adjust expenditures. Here we focus on instructional expenditures since these are the most likely to affect the quality of students' education. Table 3.6 presents the effects of caps and freezes on per-student instructional expenditures. We see a negative effect of 3.3 percentage points during a cap/freeze. This aligns with results from Bound and Turner (2007) which show that universities decrease expenditures per student when the size of a cohort is large. Additionally, results show large heterogeneity by institutional characteristics. Colleges that are *More Dependent* on tuition decrease instructional expenditure by 5.0 percentage points per year during a regulation, relative to what they would have spent in the absence of the freeze. Effects are also magnified for the Carnegie *Others* group of colleges during and after the regulation.

By further decomposing instructional expenditures into subcategories, we find that the negative effect on per-student instructional expenditures is mainly driven by universities' tightening fringe benefits for instructional staff. We see a negative effect of 4.5 percentage points on the log of total benefits for instructional staff per student. Meanwhile, we do not find evidence that universities downsize instructional staff or decrease the baseline salary, both of which may be less adjustable in the short-run than fringe benefits. Analysis of results are presented in Table 3.18. In addition, during the period of analysis, the average amount of fringe benefits is equivalent to 25% of the average salary ($15,544 and $58,657, respectively, 2011 CPI adjusted.)

These results imply that tuition regulations do not only affect tuition - they could have meaningful impacts on instructors which could in turn affect the quality of education.

3.5.6 Spillover Effects on Private Schools

We also investigate if tuition regulations have spillover effects on private colleges located within the same state. Tuition regulations do not apply to private colleges, but they may respond to tuition regulations since they are competing for students with the regulated public institutions.[34] In Table 3.7, we compare private institutions whose competing public institutions are under tuition caps/freezes to private institutions whose competitors are not regulated. Thus, $1(\mathtt{TuitReg}_t)_{it}$ is equal to one if a public university in the same state is under a tuition cap or freeze at time t.

Our results suggest a spillover effect of tuition regulations on private colleges' tuition and aid. Private colleges do not adjust the level of tuition during a tuition cap/freeze, but there are some negative effects in the post-tuition-regulation period. Meanwhile, they decrease institutional aid by 5 percentage points during tuition regulations, with a lingering effect after the regulation is lifted in a similar pattern to our main analysis. Notably, the magnitude of the coefficients are around one-third to half of the magnitude of the effects on public institutions shown in Table 3.4. Columns (3)-(8) of Table 3.7 present spillover effects by Carnegie classification. Negative effects on tuition and aid are largely driven by *Other* institutions rather than *Research* or *Non-Research* universities. This aligns with our main heterogeneity analysis in subsection 3.5.3 showing the strongest responses from public *Other* institutions and is intuitive given that private institutions are likely to compete with public institutions of similar characteristics such as selectivity or resource availability.

[34] Previous papers have studied how colleges set tuition and aid in an equilibrium framework (Epple et al., 2006; Fu, 2014). Epple et al. (2006) consider a setting where private colleges set financial aid strategically, predicting that a student would get the same aid offer from all private colleges when her academic preparedness is common knowledge among colleges. Although our setting studies private institutions' responses to decisions of public institutions while they focus on competition among private institutions, our results are in line with their prediction.

3.5.7 Other Outcomes

Student Fees, Room and Board Charges If tuition regulations do not include limits on additional student fees, we may expect to see an increase in fees during and after the regulation. However, appendix Table 3.20 shows that fees are not affected very much, aside from some suggestive evidence that two-year colleges increase fees in the first and second year after a regulation ends. It could be that effects are dampened by some states that also limit student fees in their tuition regulations (e.g. North Carolina, Ohio, Virginia). Appendix Table 3.20 also shows the effect of tuition regulations on room and board, another potential margin that colleges could adjust to make up for lost tuition revenue. However, we do not find any evidence of this behavior.

Out-of-state Student Tuition and Enrollment Appendix Table 3.21 illustrates the effect of tuition caps and freezes on out-of-state tuition and the composition of enrolled students by state residency. We restrict our sample to 4-year institutions given that 2-year institutions enroll few out-of-state students. We do not see a clear pattern of effects of tuition regulations on these outcome variables. Notably, colleges do not hike up out-of-state tuition to compensate for losses from freezing in-state tuition. Our lack of significant changes in out-of-state tuition may be related to colleges not having market power in the out-of-state student market, making them essentially price-takers. Rizzo and Ehrenberg (2004) also find that public institutions use out-of-state students to increase institutional quality, not to increase revenue.

Completion Rate We may expect the decrease in the expenditure per student and aid to impact completion rates (Dynarski, 2003; Bound and Turner, 2007; Dynarski and Scott-Clayton, 2013; Bettinger et al., 2019; Anderson, 2020). However, we do not find any strong evidence that tuition regulations impact completion rates. It could be because we can not separately identify completion rates of low-income students, who are known to benefit the most from generous financial aid (Anderson, 2020).[35] Column (1) in appendix Table 3.22 presents these results.

3.6 Representative Student's Change in Tuition Paid

So far, we have shown that tuition regulations have meaningful impacts on in-state tuition and institutional financial aid and that these impacts vary over time and across different types of colleges. However, it is difficult to see a clear picture of the overall impact that one of these regulations might have on a student moving through their education around the time of one of these regulations. In this section, we summarize the effects that tuition regulations have on several "representative" students that differ in the types of university they are attending as well as whether they receive institutional financial aid. We also incorporate differences in the dynamics of tuition and financial aid during and after a cap or freeze by presenting estimates for two types of students who start their education at different times. First, we consider a student who begins their four-year education in the first year of a tuition regulation. For simplicity, we assume that the tuition regulation lasts 3 years, which is the median length of tuition regulations in our data. Next, we consider a student who begins their four-year education in the first year after a tuition regulation has ended.

We use our estimates from appendix Table 3.11 to calculate the effect on each representative student's tuition in each year of their four-year education. This specification captures the dynamics of negative impacts on tuition increasing as the regulation lasts longer.[36] We use the average percent of tuition covered by institutional aid at four-year public institutions as a baseline for the portion of

[35] IPEDS provides separate graduation rates for Pell grants recipients, but only beginning in 2016. We do have access to completion rates by race and gender for a longer period of time, but we do not find any meaningful patterns of tuition regulation effects on these completion rate, either.

[36] For the student who starts their education in the first year the regulation is imposed, we use $1(\mathsf{TuitReg}_t)_{it} + (\mathsf{FirstYrofTuitReg}_t)_{it}$ for their first year, $1(\mathsf{TuitReg}_t)_{it}$ for their second year, $1(\mathsf{TuitReg}_t)_{it} + 1(\mathsf{LastYrofTuitReg}_t)_{it}$ for their third year, and $1(\mathsf{TuitReg}_{t-1})_{it}$ for their fourth and final year. For the student starting right after the regulation has been lifted, we use $1(\mathsf{TuitReg}_{t-k})_{it}$ for their kth year.

tuition that is affected by changes in institutional aid. This average is unconditional on receipt of institutional aid, so our results can be interpreted as the average effect across students who do and do not receive institutional aid. For each subgroup, we compute this average within institutions of that subgroup.[37] To make the tuition and aid estimates comparable, we restrict the sample to observations that have non-missing values for both tuition and institutional aid.

Figure 3.7 presents our results. The top panel represents students starting their education in the first year of a regulation and the bottom panel represents students starting in the first year after a regulation has ended. The first column shows average effects; the second and third columns show heterogeneity in effects across types of institutions outlined in subsection 3.5.3. Tuition estimates give the percentage point change in tuition paid by the representative student, aid estimates the percentage point change in tuition paid due to changes in institutional aid received, and total estimates combine these two effects. Note that positive values for the aid column do not imply that aid is increasing, they show that the decrease in aid leads to students paying more tuition.

Focusing on the upper left panel, the top line shows that the representative student starting their education in the first year of the regulation gets a 4.3 percent discount on their tuition over the four years they are enrolled. However, the second line shows that students must pay 2.9 percent more in tuition due to their decrease in institutional aid. The bottom line shows the combination of these two effects, which reveals that they get a 1.4 percent discount overall. These separate estimates emphasize the importance of considering financial aid when thinking about how beneficial tuition regulations are to students, since without considering changes in institutional aid we would have concluded that the average discount was around triple the true discount. Students who do not receive any institutional financial aid experience the full tuition discount shown in the top row, highlighting the differences in benefits from the tuition regulation between students depending on

[37] The average percent of tuition covered by aid is 23.5 percent overall, 20.6 percent for institutions more dependent on tuition, 27.5 percent for institutions less dependent on tuition, 32.0 percent for research universities, 21.2 percent for non-research universities, and 17.5 percent for other institutions.

their institutional aid receipt.

The middle panel splits these effects into institutions that are more or less dependent on tuition revenue. Finally, the right panel shows responses by broad Carnegie classification. Benefits to students vary greatly across types of institutions and their timing of entering college. We estimate that a student who starts their education in the first year of the regulation at an institution that is *Less Dependent* on tuition will receive an overall 3.9 percent discount, but a student who starts after the regulation at a *More Dependent* on tuition institution will end up paying 2.5 percent more than they would have in the absence of the regulation. Appendix Figure 3.10 shows the corresponding figure where changes in tuition and institutional aid are measured in dollars rather than percent. Results are qualitatively similar for average and tuition dependency panels, but change for the Carnegie classification panel due to differences in tuition levels between subgroups.

To illustrate how the effects of the tuition regulation vary with the timing of student entry, we further break down the yearly effects. We focus on the subgroup of colleges that are *More Dependent* on tuition, since this is where the timing of student entry leads to the most dramatic differences in total tuition paid. As shown in Figure 3.7, students who enter in the first year of the regulation receive a 0.5 percent discount, while those who enter after the regulation ends have to pay 2.5 percent more. Figure 3.11 shows that this is driven by the deep discount in the final year of the tuition regulation, which occurs in students' junior (third) year if they started with the regulation. Meanwhile, students who start after the regulation have to pay more in the last three years of their education than they would have in the absence of the regulation. This aligns with our results presented in Figure 3.4, which show that more tuition-dependent colleges begin to raise tuition while keeping institutional aid low in the years after the regulation.

The only margins that we consider in this analysis are changes in in-state tuition and institutional aid, abstracting away from other things that may be affected. First, we do not capture any changes in application or enrollment behavior induced by the tuition regulation. Second, we assume all students complete their university education in four years, which excludes any student who drops out or takes more than four years. In addition, we don't consider any changes in educational quality resulting from the regulations. We suspect that change in institutional quality would decrease the benefits students receive from tuition caps and regulations due to the decreases in per-student instructional expenditure discussed in subsection 3.5.7 and shown in Table 3.6. We do not consider these changes in benefit calculations for simplicity, but without considering them, our results may be overstating the benefits of tuition regulations for students.

3.7 Conclusion

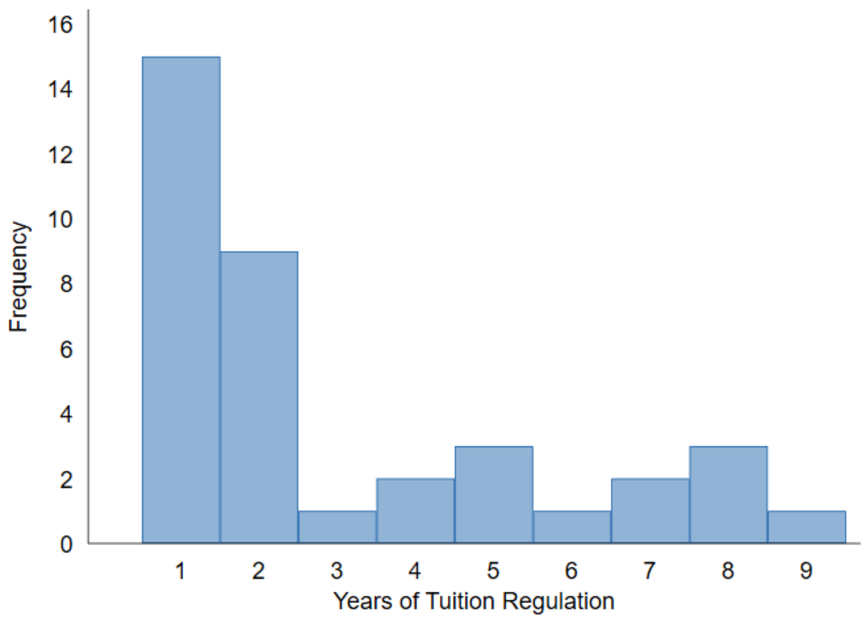
This paper has explored the effects of a popular policy tool for targeting college affordability - tuition caps and freezes. We find significant heterogeneity in the effectiveness of caps and freezes over time, with the policies only having bite in the earlier period we study (1990-2013). However, we find that when the policies have bite and tuition falls during a cap or freeze relative to where it would have been without regulations, the effects on tuition alone do not accurately reflect actual discounts for students. This is because colleges decrease their institutional financial aid when facing a tuition cap or freeze by a proportion that is almost double the decrease in tuition. Even in the years following the lifting of the regulation, institutional aid lags behind where it would have been without a regulation.

Effects of tuition regulations are not felt equally across all students. In particular, students who do not receive institutional financial aid will see much greater benefits from tuition caps and freezes than students who rely on aid. Unfortunately our institution-level data does not allow us to investigate which students see decreases in their institutional aid around the time of a tuition cap/freeze. However, we can get a sense of who is likely to be most hurt from looking at the characteristics of students who receive institutional financial aid in another data source, the National Postsecondary Student Aid Study (NPSAS).[38] Students attending four-year public colleges are more likely to receive institutional aid if they are low-income. 27

[38]Ideally we would use NPSAS data directly in our analysis, but it is only conducted every 4 years thus is not able to capture dynamics of regulations that are potentially changing every year.

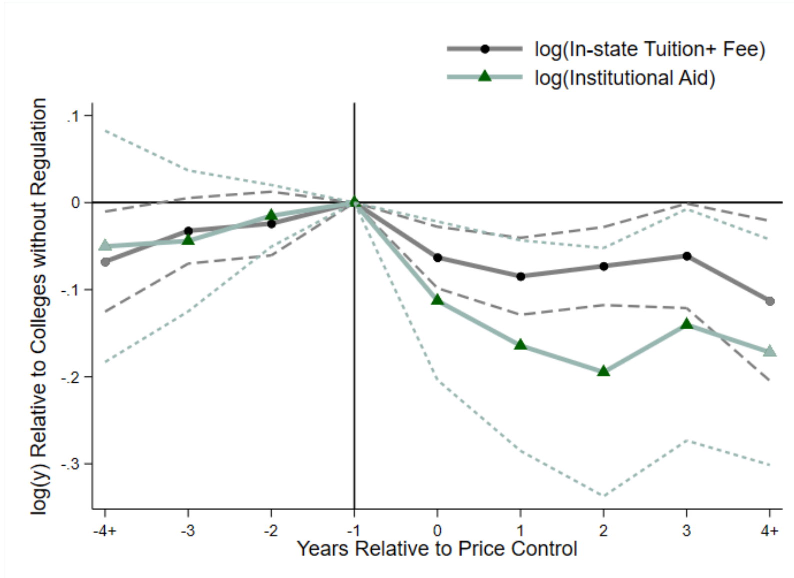
percent of students from the bottom quartile get institutional aid, as opposed to 16 percent from the top income quartile. 34 percent of students receiving Pell grants also get institutional aid, whereas only 18 percent of non-Pell-eligible students get institutional aid. This suggests that the benefit of tuition regulations may be smallest for those most in need. Further, heterogeneity analysis reveals that research institutions and institutions that do not rely heavily on tuition revenue are largely shielded from these effects, creating more inequality in how the regulations are felt by students who attend different types of colleges.

These are important impacts for policy-makers to understand. First, we have shown that tuition freezes and caps are not always effective in lowering tuition, as they have not had much bite in recent years. However, when they do have impacts on tuition, we have shown that universities respond by decreasing financial aid which disproportionately impacts students who are supported by institutional aid. This implies that tuition regulations are ineffective at best and can be harmful to needy students at worst. In the future, if policy-makers implement tuition regulations, they should be aware of these responses and consider pairing freezes/caps with policies that address the distributional consequences. Tuition freezes and caps could be accompanied with increases in financial aid or additional regulations that freeze or prohibit decreases in institutional aid.

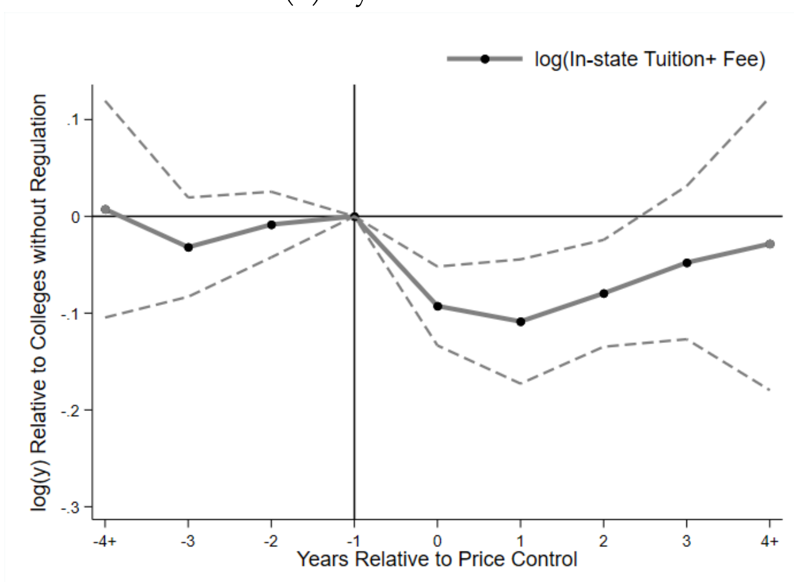


Notes: A regulation continued over multiple years is counted as one regulation. Each individual regulation represents a continuous state-level freeze or cap lasting for the specified number of years.

Figure 3.1: Distribution of Length of Tuition Regulations



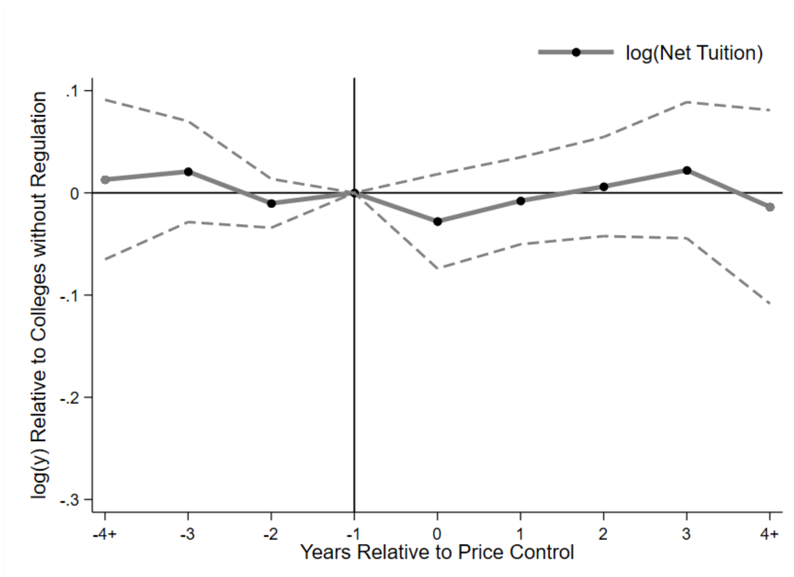
(a) 4-yr Institution



(b) 2-yr Institution

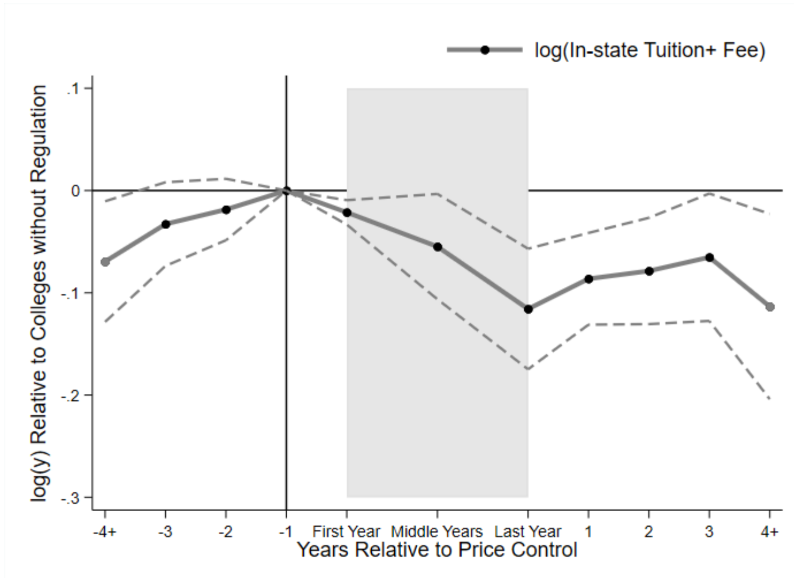
Notes: -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. The values of coefficients in the top panel are presented in Table 3.4; the bottom panel in Table 3.10. Confidence interval at 95% level.

Figure 3.2: Effect of Tuition Regulation on Tuition and Aid

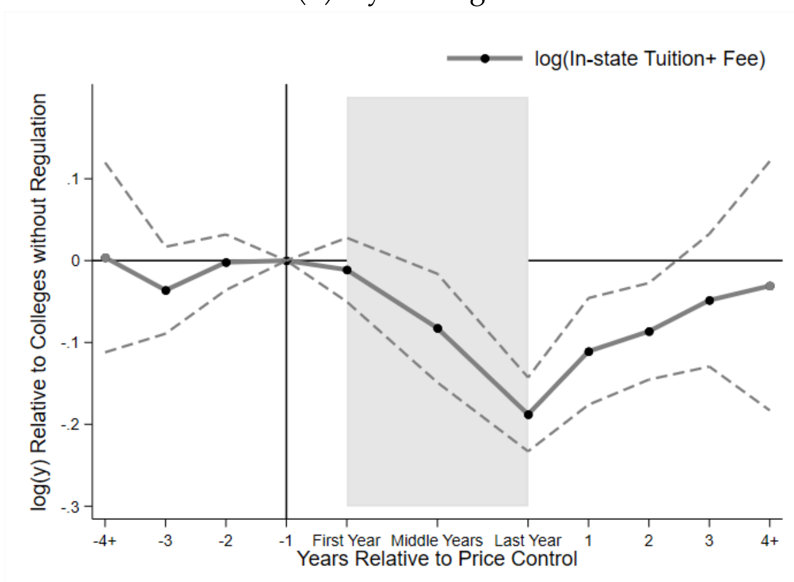


Notes: -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. Authors calculated the net tuition by subtracting the average institutional aid from tuition. Confidence interval at 95% level.

Figure 3.3: Effect of Tuition Regulation on Net Tuition



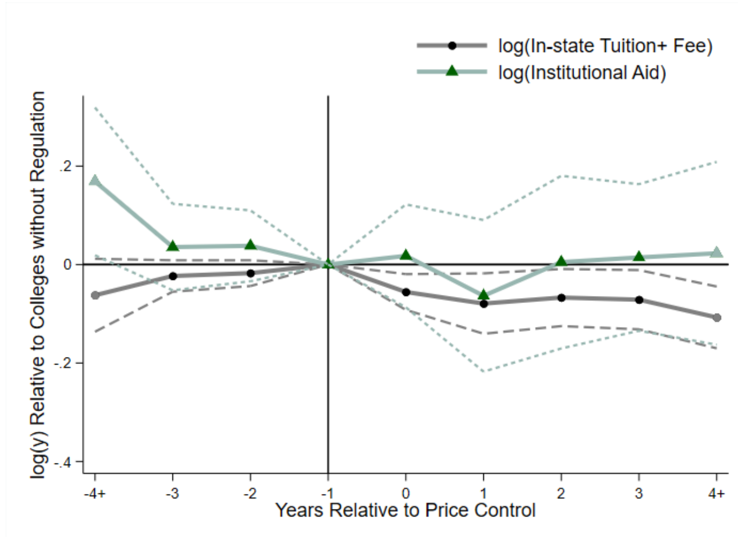
(a) 4-yr Colleges



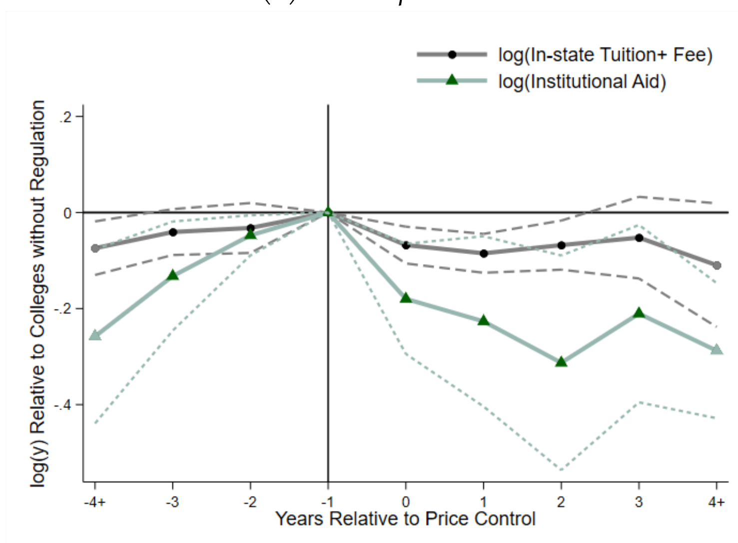
(b) 2-yr Colleges

Notes: -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. The values of coefficients are presented in Table 3.11. Confidence interval at 95% level.

Figure 3.4: Effect of Tuition Regulation on Tuition, First and Last Year of Regulation



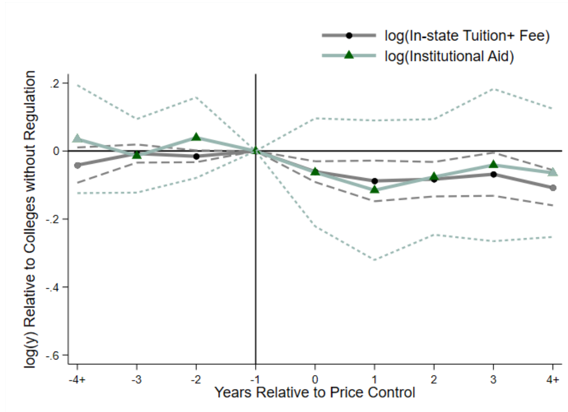
(a) *Less Dependent*



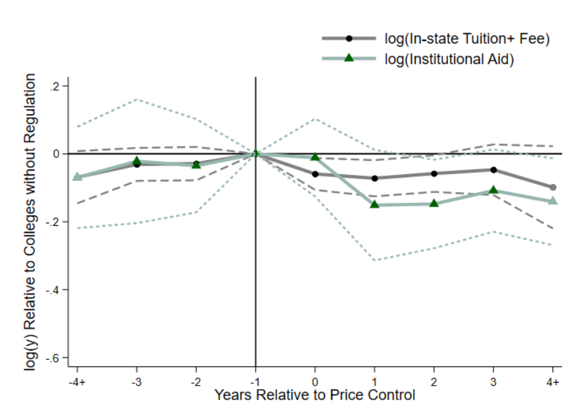
(b) *More Dependent*

Notes: Sample of 4+ year degree granting institutions. -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. The values of coefficients are presented in Table 3.13. We classify an institution into *More Dependent* if the ratio of gross tuition revenue to total revenue is above the median of the institutions in the same sector (pubic and private separately) in 1991; *Less Dependent* if below the median. Confidence interval at 95% level.

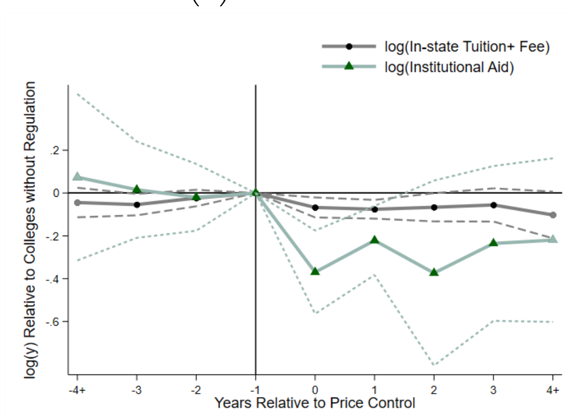
Figure 3.5: Effect of Tuition Regulation on Tuition and Aid: by Tuition Revenue Dependency



(a) *Research*



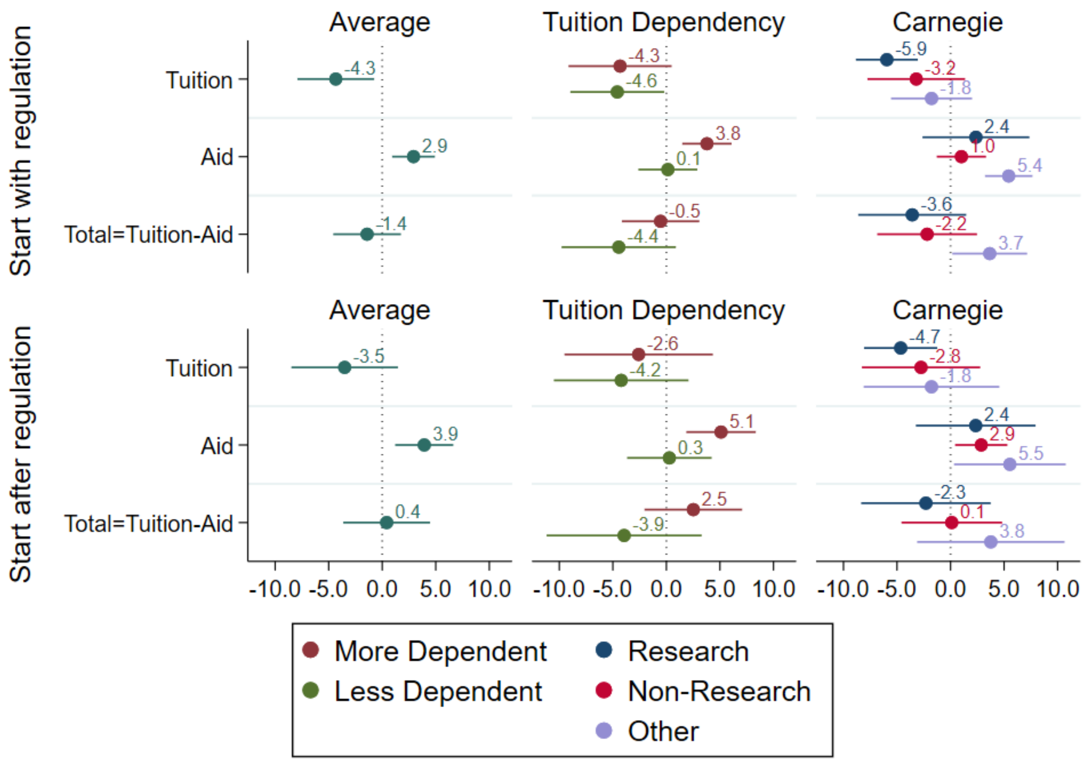
(b) *Non-research*



(c) *Others*

Notes: Sample of 4+ year degree granting institutions. -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. The values of coefficients are presented in Table 3.14. *Research* sample is of doctoral universities with high or very high research activity (Carnegie classification). *Non-Research* is sample of master's universities or Doctoral universities with low research activity. *Others* include all other 4+ year degree granting institutions. Confidence interval at 95% level.

Figure 3.6: Effect of Tuition Regulation on Tuition and Aid: by Carnegie Classification



Notes: Sample of 4+ year degree granting institutions. Tuition gives the percentage point change in net tuition paid for an average student at each type of university based on out estimates of change in listed tuition only. Aid gives the percentage point change in tuition paid due to changes in institutional aid. It is constructed by multiplying our estimates of the percent change in institutional aid with the (unconditional) percent of tuition covered by aid in each subgroup before any tuition regulations are imposed. Total combines these two effects to give the overall percentage point change in net tuition paid by a student who receives the average institutional aid, including those who receive no institutional aid. All calculations assume that the tuition regulation lasts 3 years and students attend college for 4 years. The top row gives the effect on a student whose first year of education is the first year of the regulation; the bottom row gives the effect on a student whose first year of education is the first year after the end of the regulation. Subgroups are defined as in the text. Confidence intervals at the 95% level.

Figure 3.7: Percent Change in Net Tuition Paid for Representative Students

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	Treated		Public 4-year		Private 4-year		Public 2-year		Private 2-year	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
In-state Tuition										
$, 2016 referenced	5,166.55	3,361.55	5,531.35	2,690.30	18,850.56	9,482.26	2,765.61	1,756.91	9,343.28	5,675.13
% annual growth	0.001	0.068	0.044	0.078	0.033	0.096	0.041	0.157	0.028	0.14
Out-of-state Tuition										
$, 2016 referenced	11,815.97	6,094.78	13,563.50	5,599.00	18,866.66	9,469.98	6,265.22	3,155.81	9,490.43	5,737.69
% annual growth	0.006	0.094	0.036	0.109	0.032	0.096	0.027	0.193	0.027	0.146
Average Institutional aid										
$, 2016 referenced	935.239	1,283.43	1,279.78	1,213.97	7,814.53	5,616.63	256.167	410.341	1,313.53	2,443.63
% annual growth	0.111	0.985	0.09	0.689	0.073	0.635	0.082	1.043	0.091	1.139
% Revenue Souced with Tuition	0.336	0.188	0.287	0.144	0.639	1.29	0.223	0.132	0.669	3.123
Carnegie Classification										
Others	0.32	0.466	0.247	0.431	0.603	0.489	-	-	-	-
Non-research	0.333	0.471	0.452	0.498	0.339	0.473	-	-	-	-
Research	0.347	0.476	0.301	0.459	0.058	0.234	-	-	-	-
N of Obs	2,636		13,856		29,025		23,908		4,683	
N of Aid Obs	2,012		8,761		17,903		14,440		1,842	

Notes: 1. The unit of observation is Year × Institution. 2. Variables in dollar amount are adjusted using Consumer Price Index (CPI). Deflator of 2016 is normalized to be 100. 3. Tuition is the sum of undergraduate tuition and fee. 4. *Research* sample is of doctoral universities with high or very high research activity (Carnegie classification). *Non-Research* is sample of master's universities or Doctoral universities with low research activity. *Others* include all other 4+ year degree granting institutions. 5. % Revenue soured with tuition is the fraction of gross tuition revenue out of total revenue.

Table 3.1: Summary Statistics by Type of Institution

	(1)	(2)	(3)	(4)	(5)	(6)
	All		4-year Institution		2-year Institution	
	Pre-2013	Post-2014	Pre-2013	Post-2014	Pre-2013	Post-2014
$1(\mathtt{TuitReg}_t)_{it}$	-0.112***	-0.001	-0.075***	-0.021	-0.143***	0.010
	(0.035)	(0.013)	(0.025)	(0.013)	(0.050)	(0.020)
$\mathtt{TuitCap}_{it}$	1.007**	-0.128	0.977***	-0.007	0.822	-0.196
	(0.436)	(0.214)	(0.302)	(0.239)	(0.673)	(0.250)
Observations	70,845	14,556	41,360	9,776	29,485	4,780
R-squared	0.798	0.397	0.857	0.440	0.715	0.352
Two-way FEs	yes	yes	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes	yes

Notes: 1. Pre-2013 includes 2013 and years before. Post-2014 includes 2014 and years after. 2. The outcome variables are the log of in-state undergraduate tuition and fees combined in all columns. 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. A private/public specific time trend is included. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.2: Effect of Tuition Regulation on Tuition: Time Periods Before 2013 and After 2014

	Under Tuition Regulation			Not Under Regulation		
	N	mean	sd	N	mean	sd
Before 2013	2,664	0.024	0.075	69,239	0.063	0.139
After 2014	2,019	0.025	0.048	14,724	0.031	0.070

Notes: 1. Pre-2013 includes 2013 and years before. Post-2014 includes 2014 and years after. 2. 4-year and 2-year institution pooled.

Table 3.3: Annual Tuition Increase Rate Before and After 2013

Dep. Variable	(1) log(In-state Tuition)	(2)	(3) log(Institutional Aid)	(4)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t+k})_{it}$	-0.068	-0.063	-0.050	-0.044
	(0.029)	(0.028)	(0.066)	(0.061)
$1(TuitReg_{t+3})_{it}$	-0.032	-0.032	-0.044	-0.034
	(0.019)	(0.019)	(0.040)	(0.041)
$1(TuitReg_{t+2})_{it}$	-0.024	-0.023	-0.015	-0.016
	(0.018)	(0.018)	(0.018)	(0.020)
$1(TuitReg_{t})_{it}$	-0.063	-0.094	-0.113	-0.101
	(0.018)	(0.020)	(0.045)	(0.046)
$1(TuitReg_{t-1})_{it}$	-0.085	-0.115	-0.164	-0.201
	(0.022)	(0.031)	(0.060)	(0.070)
$1(TuitReg_{t-2})_{it}$	-0.073	-0.100	-0.195	-0.280
	(0.022)	(0.028)	(0.071)	(0.091)
$1(TuitReg_{t-3})_{it}$	-0.061	-0.094	-0.140	-0.186
	(0.030)	(0.031)	(0.066)	(0.088)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t-k})_{it}$	-0.113	-0.110	-0.172	-0.162
	(0.046)	(0.044)	(0.064)	(0.060)
$TuitCap_{it}$		0.967		-0.199
		(0.322)		(0.373)
$TuitCap_{it-1}$		1.024		1.397
		(0.528)		(1.064)
$TuitCap_{it-2}$		0.831		2.837
		(0.478)		(1.606)
$TuitCap_{it-3}$		0.871		1.434
		(0.337)		(1.848)
Observations	41,410	41,410	26,239	26,239
R-squared	0.856	0.857	0.293	0.293
Two-way FEs	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. The outcome variables are the log of in-state undergraduate tuition and fees combined in columns (1)-(2), and the log of average institutional aid for first-time undergraduates in column (3)-(4). 2. Two-way fixed effects include institution fixed effects and year fixed effects. 3. A private/public specific time trend is included. 4. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 5. Standard errors clustered at the state level are in parentheses.

Table 3.4: Effect of Tuition Regulation on Tuition and Aid: 4-year Institution

Dep. Variable	(1) Gross	(2) Net	(3) log(Gross)	(4) log(Net)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it}$	-6.928	-5.126	0.048	0.015
	(8.920)	(6.716)	(0.045)	(0.059)
$1(\mathrm{TuitReg}_{t+3})_{it}$	-2.367	-1.099	0.046	0.075
	(3.035)	(2.156)	(0.024)	(0.038)
$1(\mathrm{TuitReg}_{t+2})_{it}$	-2.630	-1.262	0.024	0.060
	(1.525)	(1.009)	(0.015)	(0.032)
$1(\mathrm{TuitReg}_{t})_{it}$	-4.720	-2.675	-0.035	-0.023
	(2.615)	(2.027)	(0.028)	(0.049)
$1(\mathrm{TuitReg}_{t-1})_{it}$	-3.287	-3.461	-0.013	-0.012
	(4.854)	(2.897)	(0.030)	(0.054)
$1(\mathrm{TuitReg}_{t-2})_{it}$	-3.721	-3.828	0.009	0.033
	(4.973)	(3.039)	(0.030)	(0.047)
$1(\mathrm{TuitReg}_{t-3})_{it}$	-2.955	-3.727	0.019	0.031
	(5.603)	(3.371)	(0.029)	(0.046)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it}$	-3.220	-3.511	-0.021	-0.019
	(7.740)	(4.555)	(0.033)	(0.051)
Observations	31,944	32,050	31,943	32,048
R-squared	0.248	0.229	0.604	0.430
Two-way FEs	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. 2. The outcome variables are gross tuition revenue (in millions) in column (1), net tuition revenue (in million) in column (2), and the log of gross/net tuition revenue in column (3) and (4), respectively. 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. A private/public specific time trend is included. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.5: Effect of Tuition Regulation on Tuition Revenue

Sample	(1) All	(2) Tuition Dependency Less Dep.	(3) Tuition Dependency More Dep.	(4) Carnegie Classification Other	(5) Carnegie Classification Non-research	(6) Carnegie Classification Research
$\Sigma_{k=4}^{\infty} 1(\text{TuitReg}_{t+k})_{it}$	-0.028	-0.068	-0.005	-0.070	-0.032	0.010
	(0.029)	(0.030)	(0.027)	(0.039)	(0.031)	(0.018)
$1(\text{TuitReg}_{t+3})_{it}$	-0.021	-0.035	-0.018	-0.047	-0.023	0.004
	(0.017)	(0.020)	(0.011)	(0.034)	(0.011)	(0.011)
$1(\text{TuitReg}_{t+2})_{it}$	0.000	-0.013	0.001	-0.027	-0.004	0.006
	(0.014)	(0.019)	(0.011)	(0.028)	(0.012)	(0.007)
$1(\text{TuitReg}_{t})_{it}$	-0.033	-0.021	-0.050	-0.054	-0.033	-0.008
	(0.017)	(0.019)	(0.018)	(0.028)	(0.021)	(0.009)
$1(\text{TuitReg}_{t-1})_{it}$	-0.022	-0.040	-0.027	-0.074	-0.010	-0.019
	(0.017)	(0.018)	(0.024)	(0.031)	(0.020)	(0.015)
$1(\text{TuitReg}_{t-2})_{it}$	-0.027	-0.042	-0.029	-0.087	-0.013	-0.008
	(0.019)	(0.016)	(0.029)	(0.031)	(0.023)	(0.017)
$1(\text{TuitReg}_{t-3})_{it}$	-0.024	-0.046	-0.025	-0.080	-0.010	-0.014
	(0.024)	(0.020)	(0.027)	(0.029)	(0.024)	(0.014)
$\Sigma_{k=4}^{\infty} 1(\text{TuitReg}_{t-k})_{it}$	-0.020	-0.044	-0.022	-0.055	-0.018	-0.023
	(0.026)	(0.026)	(0.018)	(0.030)	(0.021)	(0.014)
Observations	44,694	20,347	20,485	19,392	15,443	5,463
R-squared	0.492	0.535	0.679	0.450	0.743	0.627
Two-way FEs	yes	yes	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. We classify an institution into *More Dependent* if the ratio of gross tuition revenue to total revenue is above the median of the institutions in the same sector (pubic and private separately) in 1991; *Less Dependent* if below the median. Research sample is of doctoral universities with high or very high research activity (Carnegie classification). Non-Research is sample of master's universities or Doctoral universities with low research activity. Others include all other 4+ year degree granting institutions. 2. The outcome variable is log of per-student Instruction-related Expenditure in all columns. 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. A private/public specific time trend is included. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.6: Effect of Tuition Regulation on Per-Student Instruction-Related Expenditure

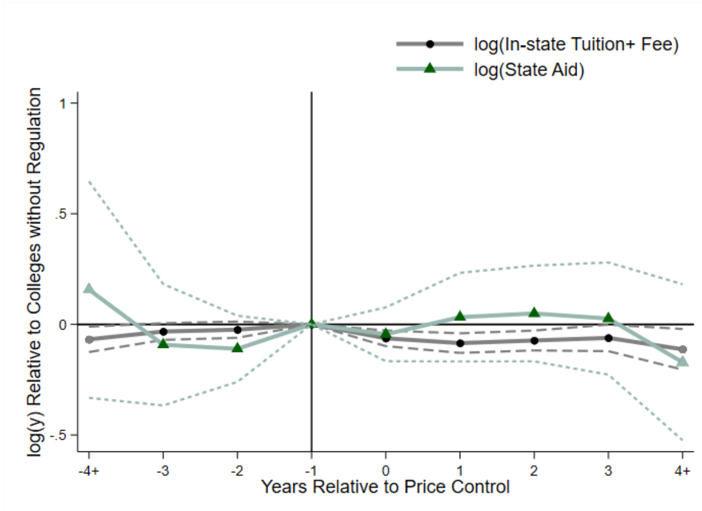
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Carnegie Classification			
Sample	All		Other		Non-research		Research	
Dep. Variable	log(Tuition)	log(Aid)	log(Tuition)	log(Aid)	log(Tuition)	log(Aid)	log(Tuition)	log(Aid)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it}$	-0.016	-0.051	-0.027	0.013	-0.004	-0.102	0.001	0.104
	(0.013)	(0.032)	(0.022)	(0.044)	(0.011)	(0.042)	(0.013)	(0.072)
$1(\mathrm{TuitReg}_{t+3})_{it}$	-0.012	-0.070	-0.018	-0.050	-0.010	-0.062	0.002	-0.045
	(0.010)	(0.040)	(0.014)	(0.062)	(0.011)	(0.044)	(0.006)	(0.095)
$1(\mathrm{TuitReg}_{t+2})_{it}$	-0.003	-0.021	-0.011	-0.010	0.002	-0.008	0.003	0.052
	(0.004)	(0.025)	(0.007)	(0.041)	(0.003)	(0.033)	(0.004)	(0.054)
$1(\mathrm{TuitReg}_{t})_{it}$	-0.004	-0.059	-0.002	-0.065	-0.001	0.001	-0.008	-0.068
	(0.005)	(0.021)	(0.006)	(0.028)	(0.005)	(0.028)	(0.005)	(0.070)
$1(\mathrm{TuitReg}_{t-1})_{it}$	-0.008	-0.088	-0.012	-0.086	-0.000	-0.029	-0.014	-0.007
	(0.008)	(0.032)	(0.008)	(0.047)	(0.010)	(0.042)	(0.009)	(0.036)
$1(\mathrm{TuitReg}_{t-2})_{it}$	-0.018	-0.099	-0.018	-0.107	-0.010	-0.023	-0.011	-0.011
	(0.009)	(0.031)	(0.009)	(0.037)	(0.012)	(0.047)	(0.009)	(0.032)
$1(\mathrm{TuitReg}_{t-3})_{it}$	-0.023	-0.107	-0.032	-0.108	-0.012	-0.020	-0.011	-0.023
	(0.010)	(0.031)	(0.013)	(0.054)	(0.013)	(0.048)	(0.009)	(0.075)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it}$	-0.027	-0.107	-0.054	-0.116	0.004	-0.036	-0.008	0.062
	(0.012)	(0.038)	(0.020)	(0.066)	(0.012)	(0.053)	(0.012)	(0.044)
Observations	30,798	18,160	14,054	8,735	10,409	6,650	1,742	1,141
R-squared	0.820	0.278	0.787	0.253	0.928	0.369	0.970	0.482
Two-way FEs	yes	yes	yes	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting private institutions. $1(\mathrm{TuitReg}_{t-k})_{it}$ equals to one if public institutions of the same state as i are under tuition regulation in $t-k$. 2. The outcome variables are the log of in-state tuition and fees combined in odd-numbered columns, and log of average institutional aid for first-time undergraduates in even-numbered columns. 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. State level controls include lag, lead and the current year of state-level unemployment rate. 5. Standard errors clustered at the state level are in parentheses.

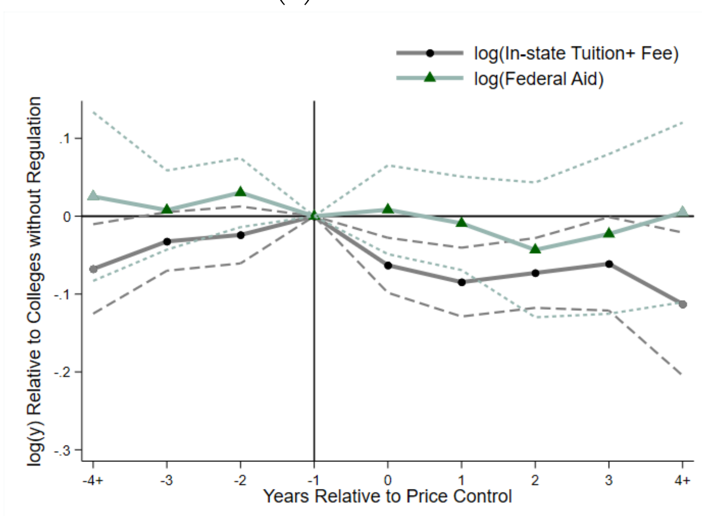
Table 3.7: Spillover Effects

3.8 Appendix

3.8.1 Supplementary Tables and Figures



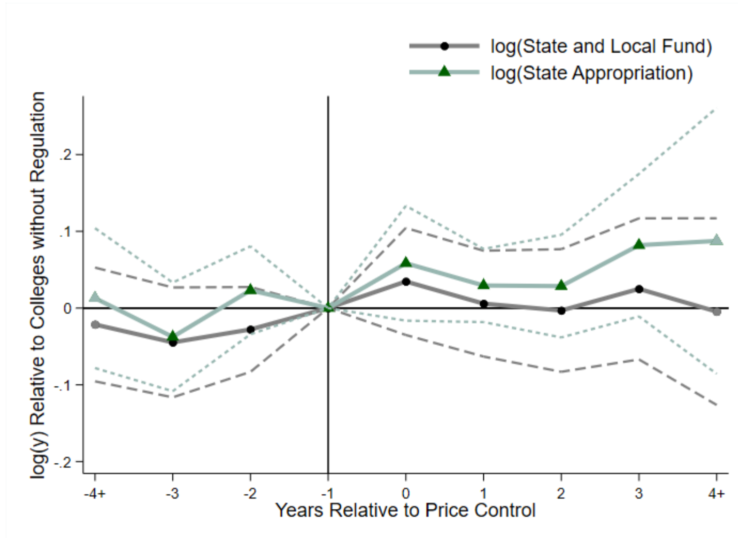
(a) State Aid



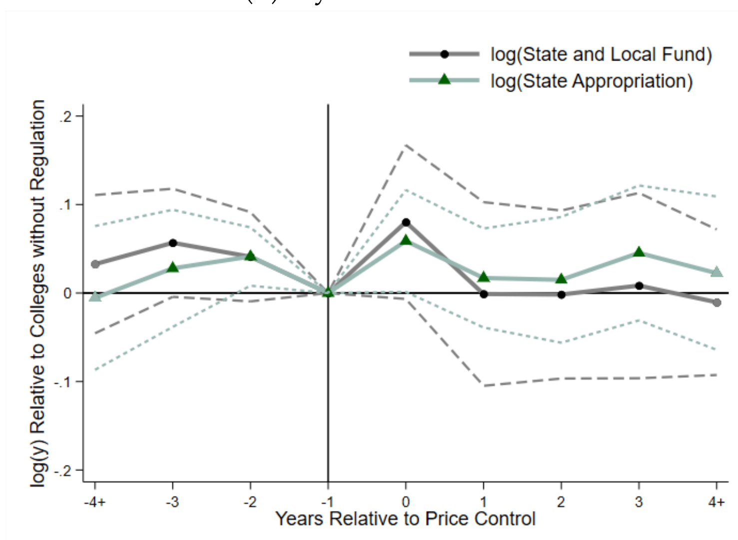
(b) Pell Grant

Notes: Sample of 4+ year degree granting institutions. -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. Confidence interval at 95% level.

Figure 3.8: Effect of Tuition Regulation on Other Sources of Aid



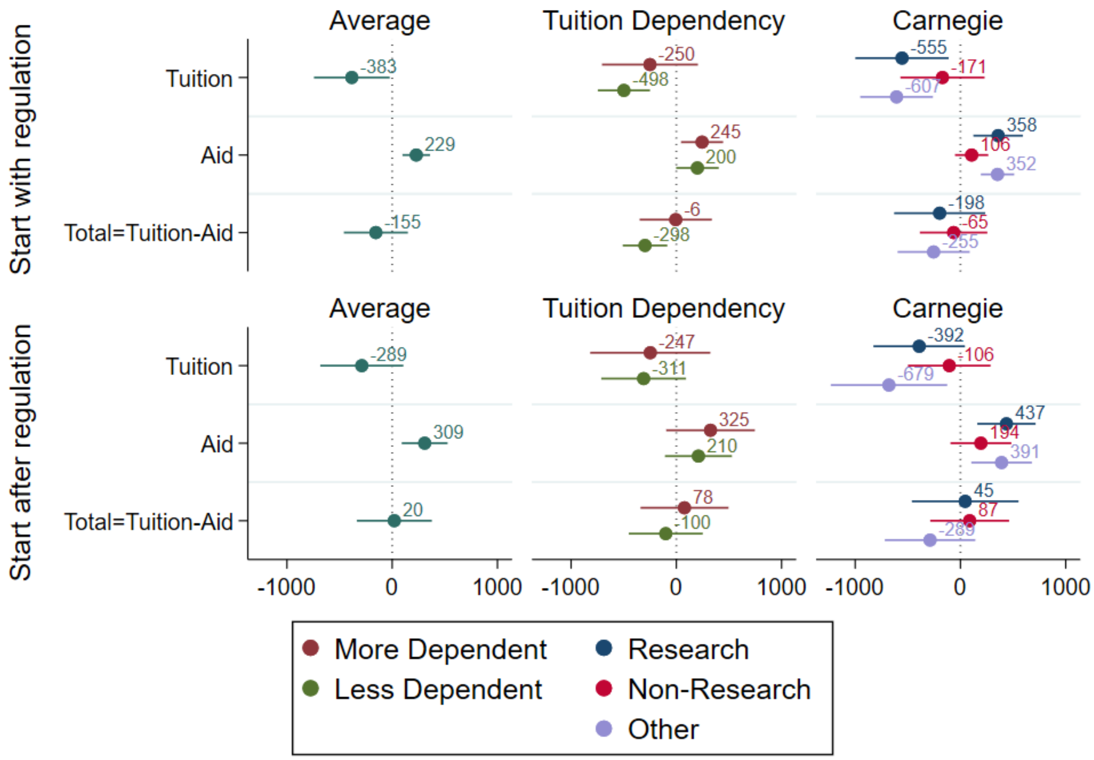
(a) 4-year Institution



(b) 2-year Institution

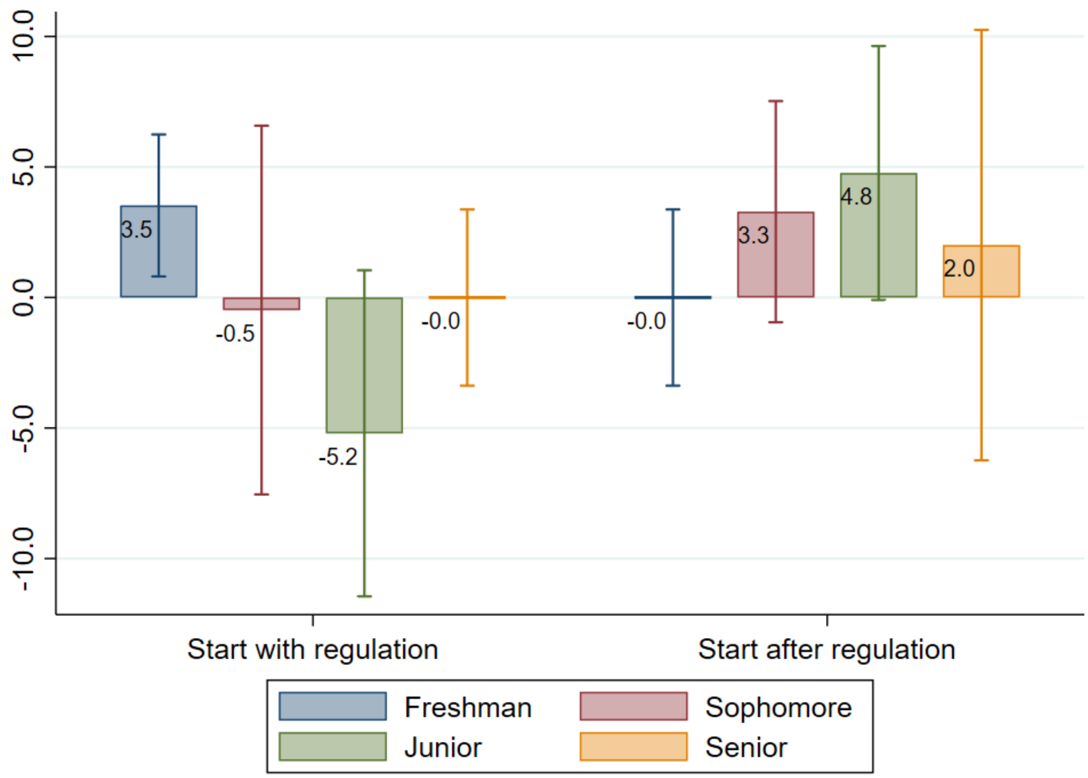
Notes: -4+ means 4 or more years before the tuition regulation is introduced, and 4+ is 4 or more years after the tuition regulation is lifted. log(State and Local Fund) is a total sum of appropriation and grants from either State or local government. log(State Appropriation) only captures the appropriation from State. Confidence interval at 95% level.

Figure 3.9: State Funding Before and After Tuition Regulation



Notes: Sample of 4+ year degree granting institutions. Tuition gives the dollar amount change in tuition paid for an average student at each type of university based on out estimates of change in listed tuition only. Aid gives the dollar amount change in tuition paid due to changes in institutional aid. Total combines these two effects to give the overall dollar amount change in tuition paid by a student who receives the average institutional aid, including those who receive no institutional aid. All calculations assume that the tuition regulation lasts 3 years and students attend college for 4 years. The top row gives the effect on a student whose first year of education is the first year of the regulation; bottom row gives the effect on a student whose first year of education is the first year after the end of the regulation. Subgroups are defined as in the text. Confidence intervals at the 95% level.

Figure 3.10: Dollar Change in Net Tuition Paid by Representative Students



Notes: Sample of 4+ year degree granting institutions. We classify an institution into *More Dependent* if the ratio of gross tuition revenue to total revenue is above the median of the institutions in the same sector (pubic and private separately) in 1991. Each year plots the total percentage point change paid in tuition incorporating changes in listed tuition and institutional aid. All calculations assume that the tuition regulation lasts 3 years and students attend college for 4 years. Left side shows results for a student whose first year of education is the first year of the regulation; right side gives results for a student whose first year of education is the first year after the end of the regulation. Confidence intervals at the 95% level.

Figure 3.11: Percent Change in Tuition Paid for Representative Students by at *More Dependent* Colleges, by Cohort

Cap	Freq.	Percent	Notes
-0.2 (mandated cut)	1	0.92	Virginia, 2000
0 (tuition freeze)	55	50.46	
0.03	8	7.34	
0.035	6	5.5	
0.04	7	6.42	
0.055	2	1.83	
0.06	12	11.01	
0.065	1	0.92	
0.07	2	1.83	
0.08	4	3.67	
0.09	1	0.92	
0.1	10	9.17	
Total	109	100	

Note: Unit of observation if year by state.

Table 3.8: Distribution of Tuition Regulations

	By State		By Year X State	
Scope	Freq.	Percent	Freq.	Percent
All public institutions	6	35.29	44	40.36
4-year public institutions	7	41.18	35	32.11
2-year public institutions	3	17.65	16	14.68
CUNY (except 2003) and Cornell	1	5.88	14	12.84
Total	17	100	109	100

Notes: Oklahoma imposed a tuition regulation on all public institutions except for Oklahoma Technology Centers. For simplicity, it is counted as in the category “All public institutions".

Table 3.9: Type of Affected Institutions

Dep. Variable	(1) log(In-state Tuition)	(2)	(3) log(Institutional Aid)	(4)
$\Sigma_{k=4}^{\infty}1(TuitReg_{t+k})_{it}$	0.007	0.004	-0.054	-0.057
	(0.056)	(0.056)	(0.084)	(0.089)
$1(TuitReg_{t+3})_{it}$	-0.032	-0.039	-0.094	-0.100
	(0.026)	(0.018)	(0.108)	(0.106)
$1(TuitReg_{t+2})_{it}$	-0.008	-0.008	-0.138	-0.154
	(0.017)	(0.016)	(0.132)	(0.140)
$1(TuitReg_{t})_{it}$	-0.093	-0.104	0.033	0.036
	(0.020)	(0.024)	(0.096)	(0.096)
$1(TuitReg_{t-1})_{it}$	-0.109	-0.130	0.087	-0.014
	(0.032)	(0.043)	(0.112)	(0.153)
$1(TuitReg_{t-2})_{it}$	-0.080	-0.086	0.089	-0.020
	(0.027)	(0.035)	(0.128)	(0.158)
$1(TuitReg_{t-3})_{it}$	-0.048	-0.056	0.233	0.176
	(0.039)	(0.036)	(0.118)	(0.133)
$\Sigma_{k=4}^{\infty}1(TuitReg_{t-k})_{it}$	-0.028	-0.035	0.159	0.175
	(0.075)	(0.081)	(0.178)	(0.165)
$TuitCap_{it}$		0.749		2.069
		(0.671)		(0.774)
$TuitCap_{it-1}$		1.235		5.900
		(1.455)		(2.359)
$TuitCap_{it-2}$		0.288		5.964
		(1.358)		(3.420)
$TuitCap_{it-3}$		0.382		3.530
		(1.347)		(2.789)
Observations	29,486	29,486	15,045	15,045
R-squared	0.715	0.715	0.173	0.174
Two-way FEs	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. The outcome variables are the log of in-state undergraduate tuition and fees combined in columns (1)-(2), and the log of average institutional aid for first-time undergraduates in column (3)-(4). 2. Two-way fixed effects include institution fixed effects and year fixed effects. 3. A private/public specific time trend is included. 4. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 5. Standard errors clustered at the state level are in parentheses.

Table 3.10: Effect of Tuition Regulation on Tuition and Aid: 2-year Institutions

	(1)	(2)	(3)	(4)
	4-year Institution		2-year Institution	
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t+k})_{it}$	-0.069	-0.062	0.004	0.020
	(0.029)	(0.030)	(0.058)	(0.059)
$1(TuitReg_{t+3})_{it}$	-0.033	-0.028	-0.036	-0.026
	(0.020)	(0.018)	(0.026)	(0.026)
$1(TuitReg_{t+2})_{it}$	-0.019	-0.022	-0.002	-0.007
	(0.015)	(0.016)	(0.017)	(0.016)
$1(FirstYrofTuitReg_{t})_{it}$	0.033		0.071	
	(0.025)		(0.025)	
$1(TuitReg_{t})_{it}$	-0.055	-0.023	-0.082	-0.029
	(0.026)	(0.009)	(0.033)	(0.029)
$NofConsecutiveYears - 1_{it}$		-0.019		-0.029
		(0.007)		(0.008)
$1(LastYrofTuitReg_{t})_{it}$	-0.061		-0.105	
	(0.012)		(0.033)	
$1(TuitReg_{t-1})_{it}$	-0.086	-0.085	-0.111	-0.108
	(0.022)	(0.023)	(0.032)	(0.033)
$1(TuitReg_{t-2})_{it}$	-0.079	-0.079	-0.086	-0.086
	(0.026)	(0.025)	(0.029)	(0.028)
$1(TuitReg_{t-3})_{it}$	-0.065	-0.066	-0.048	-0.049
	(0.031)	(0.033)	(0.040)	(0.040)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t-k})_{it}$	-0.113	-0.108	-0.031	-0.026
	(0.045)	(0.045)	(0.076)	(0.074)
Observations	41,410	41,410	29,486	29,486
R-squared	0.857	0.857	0.715	0.716
Two-way FEs	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions in columns (1)-(2), and 2+ but less than 4 year degree granting institutions in columns (3)-(4). 2. The outcome variable is the log of in-state undergraduate tuition and fees combined in columns (1)-(4). 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. A private/public specific time trend is included. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.11: Effect of Tuition Regulation on Tuition: Dynamics During Regulation

Dep. Variable	(1) In-state Tuition($)	(2)	(3) Institutional Aid($)	(4)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t+k})_{it}$	-465.3	-417.6	-380.7	-367.0
	(252.8)	(234.6)	(145.5)	(133.7)
$1(TuitReg_{t+3})_{it}$	-83.4	-56.3	-153.7	-119.2
	(121.5)	(126.0)	(75.7)	(71.7)
$1(TuitReg_{t+2})_{it}$	-118.4	-125.0	-64.3	-47.6
	(88.6)	(99.6)	(58.1)	(55.9)
$1(TuitReg_{t})_{it}$	-268.3	-326.0	-212.2	-243.4
	(121.8)	(98.3)	(63.3)	(52.5)
$1(TuitReg_{t-1})_{it}$	-243.7	-520.1	-292.0	-341.6
	(137.8)	(126.9)	(108.8)	(102.4)
$1(TuitReg_{t-2})_{it}$	-162.2	-510.6	-278.4	-439.8
	(162.8)	(184.9)	(139.5)	(91.5)
$1(TuitReg_{t-3})_{it}$	-129.0	-488.9	-221.1	-412.6
	(175.6)	(193.6)	(142.1)	(93.5)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t-k})_{it}$	-558.6	-518.7	-436.4	-419.8
	(218.8)	(198.7)	(99.0)	(100.9)
$TuitCap_{it}$		2,217.4		1,158.0
		(2,976.6)		(797.8)
$TuitCap_{it-1}$		10,135.8		1,509.2
		(4,982.4)		(1,290.7)
$TuitCap_{it-2}$		11,970.1		5,071.7
		(4,392.0)		(1,763.8)
$TuitCap_{it-3}$		11,059.2		5,610.4
		(3,642.0)		(1,906.6)
Observations	41,539	41,539	26,446	26,446
R-squared	0.819	0.819	0.612	0.612
Two-way FEs	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. 2. The outcome variables are in-state tuition and fees combined in columns (1)-(2), and the mean institutional aid in column (3)-(4). 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. A private/public specific time trend is included. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.12: Effect of Tuition Regulation on Tuition and Aid: Dollar Amount

Dep. Variable	log(In-state Tuition)		log(Institutional Aid)	
Sample	*Less Dep.*	*More Dep.*	*Less Dep.*	*More Dep.*
	(1)	(2)	(3)	(4)
$\Sigma_{k=4}^{\infty} 1(\text{TuitReg}_{t+k})_{it}$	-0.062	-0.074	0.169	-0.258
	(0.037)	(0.028)	(0.075)	(0.090)
$1(\text{TuitReg}_{t+3})_{it}$	-0.023	-0.041	0.036	-0.132
	(0.016)	(0.024)	(0.044)	(0.056)
$1(\text{TuitReg}_{t+2})_{it}$	-0.017	-0.032	0.038	-0.047
	(0.013)	(0.026)	(0.036)	(0.021)
$1(\text{TuitReg}_{t})_{it}$	-0.056	-0.068	0.018	-0.180
	(0.018)	(0.019)	(0.052)	(0.057)
$1(\text{TuitReg}_{t-1})_{it}$	-0.079	-0.085	-0.063	-0.227
	(0.031)	(0.020)	(0.077)	(0.088)
$1(\text{TuitReg}_{t-2})_{it}$	-0.067	-0.068	0.005	-0.313
	(0.029)	(0.025)	(0.087)	(0.111)
$1(\text{TuitReg}_{t-3})_{it}$	-0.071	-0.052	0.014	-0.210
	(0.030)	(0.042)	(0.074)	(0.092)
$\Sigma_{k=4}^{\infty} 1(\text{TuitReg}_{t-k})_{it}$	-0.107	-0.110	0.023	-0.288
	(0.031)	(0.064)	(0.092)	(0.070)
Observations	17,476	19,513	11,268	12,921
R-squared	0.844	0.920	0.296	0.333
Two-way FEs	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. We classify an institution into *More Dependent* if the ratio of gross tuition revenue to total revenue is above the median of the institutions in the same sector (pubic and private separately) in 1991; *Less Dependent* if below the median. 2. The outcome variables are the log of in-state undergraduate tuition and fees combined in columns (1)-(2) and the log of average institutional aid for first-time undergraduates in columns (3)-(4). 4. Two-way fixed effects include institution fixed effects and year fixed effects. 4. Column (1),(4) controls for quadratic sector-specific time trend, column (2),(5) state-specific linear time trend, and (3),(6) both sector- and state-specific linear time trend. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.13: Effect of Tuition Regulation on Tuition and Aid: by Tuition Revenue Dependency

	(1)	(2)	(3)	(4)	(5)	(6)
	log(In-state Tuition)			log(Institutional Aid)		
	Other	*Non-research*	*Research*	*Other*	*Non-research*	*Research*
$\Sigma_{k=4}^{\infty} 1(\text{TuitReg}_{t+k})_{it}$	-0.044	-0.069	-0.041	0.074	-0.069	0.035
	(0.034)	(0.038)	(0.026)	(0.193)	(0.074)	(0.079)
$1(\text{TuitReg}_{t+3})_{it}$	-0.054	-0.031	-0.007	0.015	-0.022	-0.014
	(0.025)	(0.024)	(0.013)	(0.112)	(0.091)	(0.054)
$1(\text{TuitReg}_{t+2})_{it}$	-0.023	-0.029	-0.015	-0.020	-0.035	0.039
	(0.019)	(0.024)	(0.009)	(0.078)	(0.068)	(0.059)
$1(\text{TuitReg}_{t})_{it}$	-0.067	-0.059	-0.061	-0.370	-0.011	-0.063
	(0.023)	(0.023)	(0.015)	(0.097)	(0.057)	(0.079)
$1(\text{TuitReg}_{t-1})_{it}$	-0.076	-0.072	-0.088	-0.222	-0.151	-0.115
	(0.022)	(0.026)	(0.030)	(0.080)	(0.081)	(0.102)
$1(\text{TuitReg}_{t-2})_{it}$	-0.067	-0.058	-0.083	-0.374	-0.148	-0.076
	(0.033)	(0.027)	(0.025)	(0.215)	(0.065)	(0.085)
$1(\text{TuitReg}_{t-3})_{it}$	-0.056	-0.047	-0.068	-0.235	-0.108	-0.041
	(0.039)	(0.037)	(0.032)	(0.180)	(0.060)	(0.111)
$\Sigma_{k=4}^{\infty} 1(\text{TuitReg}_{t-k})_{it}$	-0.102	-0.098	-0.108	-0.220	-0.141	-0.064
	(0.054)	(0.060)	(0.026)	(0.190)	(0.064)	(0.094)
Observations	16,988	15,434	5,444	10,688	10,272	3,716
R-squared	0.806	0.928	0.939	0.254	0.330	0.448
Two-way FEs	yes	yes	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. *Research* sample is of doctoral universities with high or very high research activity (Carnegie classification). *Non-Research* is sample of master's universities or Doctoral universities with low research activity. *Others* include all other 4+ year degree granting institutions. 2. The outcome variables are the log of in-state undergraduate tuition and fees combined in columns (1)-(3) and the log of average institutional aid for first-time undergraduates in columns (4)-(6). 4. Two-way fixed effects include institution fixed effects and year fixed effects. 4. Column (1),(4) controls for quadratic sector-specific time trend, column (2),(5) state-specific linear time trend, and (3),(6) both sector- and state-specific linear time trend. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.14: Effect of Tuition Regulation on Tuition and Aid: by Carnegie Classification

	(1)	(2)	(3)	(4)	(5)
Sample	Matching		Ever Treated		Aid Sample
Dep. Variable	log(In-state Tuition)	log(Institutional Aid)	log(In-state Tuition)	log(Institutional Aid)	log(In-state Tuition)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it}$	-0.072	-0.103	-0.096	-0.172	-0.011
	(0.025)	(0.079)	(0.026)	(0.103)	(0.028)
$1(\mathrm{TuitReg}_{t+3})_{it}$	-0.026	-0.099	-0.037	-0.090	0.003
	(0.016)	(0.046)	(0.013)	(0.047)	(0.014)
$1(\mathrm{TuitReg}_{t+2})_{it}$	-0.024	-0.020	-0.031	-0.061	-0.010
	(0.016)	(0.033)	(0.015)	(0.035)	(0.009)
$1(\mathrm{TuitReg}_{t})_{it}$	-0.046	-0.110	-0.041	-0.091	-0.044
	(0.013)	(0.043)	(0.015)	(0.043)	(0.019)
$1(\mathrm{TuitReg}_{t-1})_{it}$	-0.069	-0.173	-0.054	-0.157	-0.047
	(0.017)	(0.067)	(0.020)	(0.067)	(0.017)
$1(\mathrm{TuitReg}_{t-2})_{it}$	-0.064	-0.192	-0.051	-0.172	-0.032
	(0.019)	(0.075)	(0.021)	(0.078)	(0.023)
$1(\mathrm{TuitReg}_{t-3})_{it}$	-0.063	-0.135	-0.054	-0.105	-0.008
	(0.028)	(0.067)	(0.029)	(0.077)	(0.026)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it}$	-0.109	-0.200	-0.090	-0.123	-0.045
	(0.046)	(0.060)	(0.048)	(0.084)	(0.038)
Observations	5,947	3,851	4,138	2,785	25,517
R-squared	0.928	0.311	0.936	0.297	0.860
Two-way FEs	yes	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. Treated and comparison observations are 1-1 matched in column (1) and (2) based on the Mahalanobis distance in the annual tuition increase rate and the level of tuition from one to three years before regulation. Column (3) and (4) only include ever treated observations. Column (5) includes observations with non-missing institutional aid. 3. The outcome variables are log of in-state undergraduate tuition and fees combined in columns (1), (3), (5), and log of average institutional aid for first-time undergraduates in column (2)-(4). 4. Two-way fixed effects include institution fixed effects and year fixed effects. 5. A private/public specific time trend is included. 6. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 7. Standard errors clustered at the state level are in parentheses.

Table 3.15: Effect of Tuition Regulation on Tuition and Aid: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
		Tuition Dependency		Carnegie Classificiation		
	All	Less Dep.	More Dep.	Other	Non-research	Research
			Panel A: log(In-state Tuition)			
$1(\text{TuitReg}_t)_{it}$	-0.035	-0.04	-0.025	-0.026	-0.038	-0.036
	(0.006)	(0.014)	(0.004)	(0.009)	(0.014)	(0.005)
N	5350	2698	2329	880	2240	1580
			Panel B: log(Institutional AId)			
$1(\text{TuitReg}_t)_{it}$	-0.11	-0.019	-0.206	-0.32	-0.088	0.025
	(0.049)	(0.088)	(0075)	(0.100)	(0.067)	-0.131
N	3629	1761	1595	517	1507	1111

Notes: 1. Sample of 4+ year degree granting institutions. We classify an institution into *More Dependent* if the ratio of gross tuition revenue to total revenue is above the median of the institutions in the same sector (pubic and private separately) in 1991; *Less Dependent* if below the median. Research sample is of doctoral universities with high or very high research activity (Carnegie classification). Non-Research is sample of master's universities or Doctoral universities with low research activity. Others include all other 4+ year degree granting institutions. 2. The outcome variables are log of in-state undergraduate tuition and fees combined in panel A, and log of average institutional aid for first-time undergraduates in panel B. 3. DiD estimators proposed in De Chaisemartin and D'Haultfœuille (2020b) are calculated using the Stata package *did_multiplegt*. We compare the observations that is the first year of the first tuition control to not yet treated observations. 4. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 5. Standard errors clustered at the state level are in parentheses. Standard errors are calculated from bootstrapping with 50 set of samples.

Table 3.16: Effect of Tuition Regulation on Tuition and Aid: DiD estimator from De Chaisemartin and D'Haultfœuille (2020b)

Dep. Variable	(1)	(2)	(3)	(4)	(5)	(6)
	log(In-state Tuition)			log(Institutional Aid)		
	Sector-Year FE	State	State, Sector	Sector-Year FE	State	State, Sector
$\Sigma_{k=4}^{\infty}1(\text{TuitReg}_{t+k})_{it}$	-0.069	-0.100	-0.056	-0.049	-0.120	-0.037
	(0.028)	(0.025)	(0.027)	(0.069)	(0.064)	(0.069)
$1(\text{TuitReg}_{t+3})_{it}$	-0.024	-0.038	-0.025	-0.042	-0.071	-0.041
	(0.016)	(0.016)	(0.017)	(0.042)	(0.041)	(0.044)
$1(\text{TuitReg}_{t+2})_{it}$	-0.023	-0.030	-0.024	-0.002	-0.029	-0.017
	(0.017)	(0.018)	(0.018)	(0.025)	(0.017)	(0.018)
$1(\text{TuitReg}_{t})_{it}$	-0.057	-0.040	-0.059	-0.111	-0.060	-0.101
	(0.013)	(0.016)	(0.014)	(0.047)	(0.043)	(0.047)
$1(\text{TuitReg}_{t-1})_{it}$	-0.076	-0.059	-0.081	-0.154	-0.106	-0.156
	(0.017)	(0.017)	(0.019)	(0.064)	(0.055)	(0.063)
$1(\text{TuitReg}_{t-2})_{it}$	-0.075	-0.047	-0.069	-0.180	-0.129	-0.185
	(0.018)	(0.020)	(0.019)	(0.076)	(0.068)	(0.077)
$1(\text{TuitReg}_{t-3})_{it}$	-0.071	-0.024	-0.048	-0.124	-0.075	-0.137
	(0.026)	(0.027)	(0.025)	(0.069)	(0.066)	(0.075)
$\Sigma_{k=4}^{\infty}1(\text{TuitReg}_{t-k})_{it}$	-0.124	-0.049	-0.082	-0.168	-0.073	-0.158
	(0.041)	(0.044)	(0.039)	(0.066)	(0.062)	(0.077)
Observations	41,410	41,410	41,410	26,239	26,239	26,239
R-squared	0.857	0.863	0.866	0.293	0.300	0.302
Two-way FEs	yes	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. 2. The outcome variables are the log of in-state undergraduate tuition and fees combined in columns (1)-(3) and the log of average institutional aid for first-time undergraduates in columns (4)-(6). 3. Two-way fixed effects include institution fixed effects and year fixed effects. 4. Columns (1),(4) include sector-year fixed effects instead of a sector-specific linear time trend, columns (2),(5) a state-specific linear time trend, and (3),(6) both sector- and state-specific linear time trend. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.17: Effect of Tuition Regulation on Tuition and Aid: Different Time Trend

	(1)	(2)	(3)	(4)	(5)
Dep. Variable	log(Benefit Per Student)			log(Salary Per Student)	N Per Student
	Tuition Dependency				
Sample	All	Less Dep.	More Dep.	All	All
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t+k})_{it}$	-0.025	-0.058	0.004	-0.095	-0.043
	(0.040)	(0.064)	(0.052)	(0.034)	(0.016)
$1(TuitReg_{t+3})_{it}$	-0.024	-0.018	-0.018	-0.059	-0.011
	(0.026)	(0.030)	(0.050)	(0.050)	(0.008)
$1(TuitReg_{t+2})_{it}$	-0.024	-0.021	-0.017	-0.038	-0.022
	(0.020)	(0.026)	(0.033)	(0.042)	(0.005)
$1(TuitReg_{t})_{it}$	-0.045	-0.027	-0.057	-0.019	-0.002
	(0.020)	(0.022)	(0.021)	(0.029)	(0.010)
$1(TuitReg_{t-1})_{it}$	-0.034	-0.026	0.004	-0.026	-0.024
	(0.033)	(0.037)	(0.035)	(0.039)	(0.014)
$1(TuitReg_{t-2})_{it}$	-0.034	0.002	-0.019	-0.041	-0.028
	(0.034)	(0.043)	(0.033)	(0.045)	(0.018)
$1(TuitReg_{t-3})_{it}$	-0.031	-0.018	-0.031	-0.021	-0.013
	(0.028)	(0.044)	(0.031)	(0.042)	(0.026)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t-k})_{it}$	-0.060	-0.104	-0.032	-0.058	-0.014
	(0.051)	(0.122)	(0.030)	(0.041)	(0.020)
Constant	-95.971	-97.221	-98.071	-79.551	-0.769
	(7.416)	(9.495)	(8.299)	(3.017)	(2.325)
Observations	42,604	19,496	19,746	38,527	42,138
R-squared	0.400	0.432	0.578	0.254	0.008
Two-way FEs	yes	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes	yes
State level control	yes	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. 2. The outcome variable is the log of total instructional staff benefits per-student in columns (1)-(3), the log of total instructional staff salaries per-student in column (4), and the number of instructional staff per-student in column (5). 3. We classify an institution into *More Dependent* if the ratio of gross tuition revenue to total revenue is above the median of the institutions in the same sector (pubic and private separately) in 1991; *Less Dependent* if below the median. 4. Two-way fixed effects include institution fixed effects and year fixed effects. 5. A private/public specific time trend is included. 6. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 7. Standard errors clustered at the state level are in parentheses.

Table 3.18: Effect of Tuition Regulation on Instructional Staff Salary, Benefit, and Size

Dep. Variable	(1) log(In-state Tuition)	(2) log(Institutional Aid)	(3) log(In-state Tuition)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t+k})_{it}$	-0.071	-0.052	-0.035
	(0.031)	(0.073)	(0.054)
$1(TuitReg_{t+3})_{it}$	-0.032	-0.041	-0.039
	(0.019)	(0.041)	(0.026)
$1(TuitReg_{t+2})_{it}$	-0.025	-0.005	-0.025
	(0.019)	(0.021)	(0.015)
$1(TuitReg_{t})_{it}$	-0.064	-0.114	-0.089
	(0.016)	(0.046)	(0.020)
$1(TuitReg_{t-1})_{it}$	-0.081	-0.153	-0.095
	(0.022)	(0.062)	(0.035)
$1(TuitReg_{t-2})_{it}$	-0.068	-0.171	-0.053
	(0.024)	(0.080)	(0.022)
$1(TuitReg_{t-3})_{it}$	-0.073	-0.154	-0.074
	(0.030)	(0.068)	(0.047)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t-k})_{it}$	-0.123	-0.177	-0.090
	(0.050)	(0.063)	(0.070)
$log(StateLocalFund)_{t}$	-0.001	-0.003	-0.009
	(0.002)	(0.008)	(0.005)
$log(StateLocalFund)_{t-1}$	0.000	0.008	-0.015
	(0.002)	(0.010)	(0.007)
$log(StateLocalFund)_{t+1}$	0.002	0.013	0.004
	(0.002)	(0.011)	(0.011)
Observations	24,938	15,787	20,215
R-squared	0.894	0.295	0.750
Two-way FEs	yes	yes	yes
Sector specific trend	yes	yes	yes
State level control	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions in columns (1)-(2), and 2+ but less than 4 year degree granting institutions in columns (3). 2. The outcome variable is the log of in-state undergraduate tuition and fees combined in columns (1), (3) and the log of average institutional aid for first-time undergraduates in column (2). 3. log(State Local Fund) is a total sum of appropriation and grants from either State or local government. 4. Two-way fixed effects include institution fixed effects and year fixed effects. 5. A private/public specific time trend is included. 6. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 7. Standard errors clustered at the state level are in parentheses.

Table 3.19: Effect of Tuition Regulation on Tuition and Aid: Control for State Funding

Dep. Variable	(1) 4-year Institution Fee	(2) log(room and board)	(3) 2-year Institution Fee	(4) log(room and board)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t+k})_{it}$	-9.928	0.002	-32.521	-0.028
	(107.154)	(0.010)	(33.505)	(0.016)
$1(\mathrm{TuitReg}_{t+3})_{it}$	2.817	0.000	-26.917	-0.063
	(46.923)	(0.005)	(19.901)	(0.062)
$1(\mathrm{TuitReg}_{t+2})_{it}$	8.987	0.002	12.988	-0.037
	(31.422)	(0.003)	(10.185)	(0.031)
$1(\mathrm{TuitReg}_{t})_{it}$	-32.644	-0.016	41.516	-0.008
	(69.297)	(0.007)	(41.036)	(0.014)
$1(\mathrm{TuitReg}_{t-1})_{it}$	77.262	-0.012	52.511	0.028
	(102.909)	(0.011)	(27.217)	(0.018)
$1(\mathrm{TuitReg}_{t-2})_{it}$	77.306	-0.005	56.821	0.009
	(118.445)	(0.008)	(32.155)	(0.016)
$1(\mathrm{TuitReg}_{t-3})_{it}$	92.980	-0.010	51.676	0.039
	(111.926)	(0.012)	(34.774)	(0.034)
$\Sigma_{k=4}^{\infty} 1(\mathrm{TuitReg}_{t-k})_{it}$	84.222	-0.008	-39.367	0.113
	(112.792)	(0.013)	(34.892)	(0.026)
Observations	26,548	33,937	17,031	5,039
R-squared	0.173	0.863	0.158	0.709
Two-way FEs	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. The outcome variables are the undergrad in-state Fee in columns (1), (3) and log of room and board charged in columns (2), (4). 2. Two-way fixed effects include institution fixed effects and year fixed effects. 3. A private/public specific time trend is included. 4. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 5. Standard errors clustered at the state level are in parentheses.

Table 3.20: Effect of Tuition Regulation on Other Charges

Dep. Variable	(1) log(Out-of-state Tuition)	(2) % In-state Freshmen	(3) N In-state Freshmen
$\Sigma_{k=4}^{\infty}1(\mathrm{TuitReg}_{t+k})_{it}$	-0.065	0.017	50.978
	(0.031)	(0.006)	(32.781)
$1(\mathrm{TuitReg}_{t+3})_{it}$	-0.026	0.010	19.226
	(0.018)	(0.006)	(22.256)
$1(\mathrm{TuitReg}_{t+2})_{it}$	-0.011	0.010	11.905
	(0.014)	(0.006)	(20.087)
$1(\mathrm{TuitReg}_{t})_{it}$	-0.045	0.009	6.238
	(0.020)	(0.006)	(23.378)
$1(\mathrm{TuitReg}_{t-1})_{it}$	0.010	0.014	32.203
	(0.027)	(0.007)	(28.278)
$1(\mathrm{TuitReg}_{t-2})_{it}$	0.015	0.013	50.068
	(0.033)	(0.009)	(40.397)
$1(\mathrm{TuitReg}_{t-3})_{it}$	0.004	0.009	62.738
	(0.030)	(0.009)	(47.208)
$\Sigma_{k=4}^{\infty}1(\mathrm{TuitReg}_{t-k})_{it}$	-0.078	0.006	18.413
	(0.031)	(0.008)	(46.116)
Observations	41,410	8,147	8,147
R-squared	0.838	0.008	0.104
Two-way FEs	yes	yes	yes
Sector specific trend	yes	yes	yes
State level control	yes	yes	yes

Notes: 1. The outcome variables are the log of undergrad out-of-state tuition and fee combined in column (1), percentage/the number of students in fall cohort who paying in-state tuition rates in column (2) and (3), respectively. 2. Two-way fixed effects include institution fixed effects and year fixed effects. 3. A private/public specific time trend is included. 4. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 5. Standard errors clustered at the state level are in parentheses.

Table 3.21: Effect of Tuition Regulation on Out-of-state Students

Dep. Variable	(1) 150% time grad. rate	(2) SAT 75	(3) SAT 25	(4) % submitting SAT scores
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t+k})_{it}$	0.006	-0.601	0.657	0.297
	(0.005)	(3.108)	(3.468)	(1.307)
$1(TuitReg_{t+3})_{it}$	0.003	-1.293	2.626	0.432
	(0.003)	(2.025)	(2.856)	(0.711)
$1(TuitReg_{t+2})_{it}$	0.002	-0.974	-0.064	0.778
	(0.002)	(1.715)	(2.330)	(0.680)
$1(TuitReg_{t})_{it}$	-0.002	0.591	-1.337	0.339
	(0.003)	(1.870)	(2.400)	(0.680)
$1(TuitReg_{t-1})_{it}$	0.003	0.553	-1.194	-4.278
	(0.005)	(2.091)	(3.099)	(2.795)
$1(TuitReg_{t-2})_{it}$	0.003	0.812	-0.041	-0.810
	(0.006)	(2.747)	(4.184)	(1.086)
$1(TuitReg_{t-3})_{it}$	0.008	2.703	0.602	-1.460
	(0.005)	(3.501)	(4.840)	(1.209)
$\Sigma_{k=4}^{\infty} 1(TuitReg_{t-k})_{it}$	0.013	3.972	3.485	-2.988
	(0.006)	(3.929)	(5.414)	(1.334)
Observations	36,666	15,438	15,441	17,047
R-squared	0.065	0.020	0.015	0.081
Two-way FEs	yes	yes	yes	yes
Sector specific trend	yes	yes	yes	yes
State level control	yes	yes	yes	yes

Notes: 1. Sample of 4+ year degree granting institutions. 2. The outcome variables are 150% time graduation rate (=6 years) of cohort started with tuition regulation, 75 percentile of admitted students' SAT score, 25 percentile of admitted students' SAT score, and the percent of applicants submitted SAT score. 3. A private/public specific time trend is included. 4. Two-way fixed effects include institution fixed effects and year fixed effects. 5. State level controls include lag, lead and the current year of state-level unemployment rate. Two dummy variables - one if the majority of both Upper and Lower house are taken by Republicans and the other if by Democrats - are also included. 6. Standard errors clustered at the state level are in parentheses.

Table 3.22: Effect of Tuition Regulation on Graduation Rate and SAT Score

3.8.2 Tuition Freezes/Caps 1990-2019

Tuition Caps and Freezes, 1990-2007

State	Type	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Alabama	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0	0
Alaska	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Arizona	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Arkansas	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
California	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
California	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Colorado	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Connecticut	1	–	–	–	–	–	–	–	–	–	–	0	–	–	–	–	–	–	–
Delaware	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Florida	0	–	–	–	–	–	0	–	–	–	–	–	–	–	–	–	–	–	–
Georgia	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Hawaii	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Idaho	0	–	–	–	–	–	–	–	–	–	–	–	–	–	0.1	0.1	0.1	0.1	0.1
Illinois	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Indiana	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Iowa	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Kansas	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Kentucky	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Louisiana	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Maine	2	–	–	–	–	–	–	–	–	–	–	0	0	0	0	0	0	–	–
Maine	3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Maryland	1	–	–	–	–	–	–	–	–	0.04	0.04	0.04	0.04	0.04	–	–	–	–	0
Maryland	4	–	–	–	–	–	–	–	–	0.04	0.04	0.04	0.04	0.04	–	–	–	–	0
Massachusetts	5	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Michigan	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Minnesota	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Minnesota	7	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Minnesota	6	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Mississippi	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Missouri	8	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Montana	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Nebraska	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Nevada	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
New Hampshire	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0
New Jersey	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.09	0.08	0.08	0.08
New Mexico	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
New York	9	–	–	–	–	0	0	–	0	0	0	0	0	0	0	–	0	0	0
New York	10	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
North Carolina	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0	0.065
North Dakota	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
North Dakota	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Ohio	2	–	–	–	–	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	–	–	–	0.06	0.06
Ohio	11	–	–	–	–	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	–	–	–	0.06	0.06
Ohio	12	–	–	–	–	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	–	–	–	0.06	0.06
Oklahoma	2	–	–	–	–	–	–	–	–	–	–	–	–	0.07	0.07	–	–	–	–
Oklahoma	13	–	–	–	–	–	–	–	–	–	–	–	–	0.07	0.07	–	–	–	–
Oklahoma	14	–	–	–	–	–	–	–	–	–	–	–	–	0.07	0.07	–	–	–	–
Oregon	1	–	–	–	–	–	–	–	–	0	0	0	0	–	–	–	–	0.03	0.03
Pennsylvania	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Rhode Island	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Rhode Island	15	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
South Carolina	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
South Dakota	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
South Dakota	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tennessee	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Texas	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Utah	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Vermont	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Virginia	0	–	–	–	–	–	0.03	0.03	0	0	0	-0.2	0	0	–	–	–	–	–
Washington	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Washington	11	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Washington	12	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
West Virginia	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wisconsin	1	–	–	–	–	–	–	–	–	–	–	–	0	–	0.08	–	–	–	–
Wyoming	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Notes: This table lists states and years where state legislatures impose in-state tuition caps and freezes at public institutions. 1990-2013 data from Deming and Walters (2018). We collected 2014-2019 data through a combination of Lexis-Nexis searches of legislation and news articles, communication with state boards of education and legislatures, and verification using legislative records from state websites. Codes for Type: 99 means that the tuition is set by legislature. We do not include this case in the analysis. 1 - Applies only to four-year institutions in the state. 2 - Applies only to two-year institutions in the state. 3- Applies only to University of Maine System. 4- Applies only to St. Mary's college of Maryland. 5- Applies only to University of Massachusetts, Amherst. 6- Applies only to University of Minnesota System. 7- Applies to four-year Minnesota State System. 8- Applies only to four-year institutions whose tuition is above the average. 9- Applies only to CUNY (except 2003) and Cornell (all years). 10- Applies only to SUNY. 11- Applies only to State University. 12- Applies to regional campuses. 13- Applies only to Oklahoma research universities. 14- Applies only to Oklahoma regional institutions. 15 - Applies only to Rhode Island College and The Community College of Rhode Island.

Tuition Caps and Freezes, 2008-2019

State	Type	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Alabama	2	0	0	0	–	–	–	–	–	–	–	–	–
Alaska	0	–	–	–	–	–	–	–	–	–	–	–	–
Arizona	0	–	–	–	–	–	–	–	–	–	–	–	–
Arkansas	0	–	–	–	–	–	–	–	–	–	–	–	–
California	1	–	–	–	–	–	–				–		
California	2	–	–	–	–	–	–	99	99	99	99	99	99
Colorado	0	–	–	–	–	–	–	0.09	0.06	0.06	0.07	0.065	0
Connecticut	1	–	–	–	–	0	–	–	–	–	–	–	–
Delaware	0	–	–	–	–	–	–	–	–	–	–	–	–
Florida	0	0	–	–	–	–	–	–	–	–	–	–	–
Georgia	0	–	–	–	–	–	–	–	–	–	–	–	–
Hawaii	0	–	–	–	–	–	–	–	–	–	–	–	–
Idaho	0	0.1	0.1	0.1	–	0.1	0.1	–	–	–	–	–	–
Illinois	0	–	–	–	–	–	–	–	–	–	–	–	–
Indiana	0	–	–	–	–	–	–	–	–	–	–	–	–
Iowa	0	–	–	–	–	–	–	–	–	–	–	–	–
Kansas	0	–	–	–	–	–	–	–	–	0.036	–	–	–
Kentucky	0	–	–	–	–	–	–	–	–	–	–	–	–
Louisiana	0	–	–	–	–	–	–	–	–	–	–	–	–
Maine	2	–	–	–	0	–	0	–	–	–	–	–	–
Maine	3	–	–	–	–	–	–	–	–	–	–	–	–
Maryland	1	0	0	0	0.03	0.03	0.03	–	–	–	0.03	0.03	0.03
Maryland	4	0	0	0	0.03	0.03	0.03	0	–	–	0.03	0.03	0.03
Massachusetts	5	–	–	–	–	–	–	–	–	–	–	–	–
Michigan	1	–	–	–	–	0.071	0.04	0.0375	0.032	0.032	0.042	max(0.038, $475)	max(0.038, $490)
Minnesota	2	–	–	–	–	–	–	0	0	0	-0.01	0.01	0
Minnesota	7	–	–	–	–	–	–	0	0	–	0	–	0
Minnesota	6	–	–	–	–	–	–	0	0	–	–	–	0
Mississippi	0	–	–	–	–	–	–	–	–	–	–	–	–

Missouri	8	–	–	0	0	–	–	0.017	0.015	0.008	0.007	0.021	0.021
Montana	0	0	0	–	–	–	–	0	0	0	0	0	0
Nebraska	0	–	–	–	–	–	–		–	–	–	–	–
Nevada	0	–	–	–	–	–	–	–	–	–	–	–	–
New Hampshire	2	–	0	–	–	–	0	-0.05	0	0	0	0.025	–
New Jersey	1	–	–	0.03	0.04	–	–	–	–	–	–	–	–
New Mexico	0	–	–	–	–	–	–	–	–	–	–	–	–
New York	9	0	0	–	–	–	–	$300	$300	$300	0	$200	$200
New York	10	–	–	–	–	–	–	$300	$300	$300	0	$200	$200
North Carolina	1	0.065	0.065	0.065	0.065	0.065	0.065				–	–	–
North Dakota	1	–	–	–	–	–	–	–	–	0.025	0.025	0.04	0.04
North Dakota	2	–	–	–	–	–	–	–	–	0.025	0.025	0.04	0.04
Ohio	2	0.035	0.035	0.035	0.035	0.035	0.035	$100	$100	0	0	$10/credit hour	$10/credit hour
Ohio	11	0.035	0.035	0.035	0.035	0.035	0.035	max(0.02, $188)	max(0.02, $188)	0	0	0	0
Ohio	12	0.035	0.035	0.035	0.035	0.035	0.035	max(0.02, $114)	max(0.02, $114)	0	0	0	0
Oklahoma	2	–	–	0	–	–	–	0.05	0.06	0.047	0.086	0.071	0.038
Oklahoma	13	–	–	0	–	–	–	0	0.024	0.047	0.07	0.05	0.016
Oklahoma	14	–	–	0	–	–	–	0.057	0.056	0.049	0.086	0.041	0.046
Oregon	1	–	–	–	–	–	–	0.05	0.05	0.05	0.03	–	0.05
Pennsylvania	0	–	–	–	–	–	–	–	–	–	–	–	–
Rhode Island	1	–	–	–	–	–	–	0	0	–	–	–	–
Rhode Island	15	–	–	–	–	–	–	0	0	0	–	–	–
South Carolina	0	–	–	–	–	–	–	–	–	–	–	–	–
South Dakota	1	–	–	–	–	–	–	–	–	–	–	–	–
South Dakota	2	–	–	–	–	–	–	–	–	–	–	–	–
Tennessee	1	–	–	–	–	–	–	–	–	–	–	–	–
Texas	0	–	–	–	–	–	–	–	–	–	–	–	–
Utah	0	–	–	–	–	–	–	–	–	–	–	–	–
Vermont	0	–	–	–	–	–	–	–	–	–	–	–	–
Virginia	0	0.06	0.04	–	–	–	–	–	–	–	–	–	–
Washington	2	–	–	–	–	–	–	0	–	-0.05	-0.05	0.022	–

Washington	11	–	–	–	–	–	–	0	–	-0.05	-0.15	0.022	–
Washington	12	–	–	–	–	–	–	0	–	-0.05	-0.2	0.022	–
West Virginia	0	–	–	–	–	–	–	–	–	–	–	–	–
Wisconsin	1	–	–	–	–	0.055	0.055	0	0	0	0	0	0
Wyoming	0	–	–	–	–	–	–	–	–	–	–	–	–

Notes: This table lists states and years where state legislatures impose in-state tuition caps and freezes at public institutions. 1990-2013 data from Deming and Walters (2018). We collected 2014-2019 data through a combination of Lexis-Nexis searches of legislation and news articles, communication with state boards of education and legislatures, and verification using legislative records from state websites. Codes for Type: 99 means that the tuition is set by legislature. We do not include this case in the analysis. 1 - Applies only to four-year institutions in the state. 2 - Applies only to two-year institutions in the state. 3- Applies only to University of Maine System. 4- Applies only to St. Mary's college of Maryland. 5- Applies only to University of Massachusetts, Amherst. 6- Applies only to University of Minnesota System. 7- Applies to four-year Minnesota State System. 8- Applies only to four-year institutions whose tuition is above the average. 9- Applies only to CUNY (except 2003) and Cornell (all years). 10- Applies only to SUNY. 11- Applies only to State University. 12- Applies to regional campuses. 13- Applies only to Oklahoma research universities. 14- Applies only to Oklahoma regional institutions. 15 - Applies only to Rhode Island College and The Community College of Rhode Island.

REFERENCES

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak. 2017. The welfare effects of coordinated assignment: Evidence from the new york city high school match. *American Economic Review* 107(12):3635–89.

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak. 2017a. The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match. *American Economic Review* 107(12):3635–89.

Abdulkadiroğlu, Atila, Joshua D Angrist, Yusuke Narita, and Parag A Pathak. 2017b. Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica* 85(5):1373–1432.

———. 2021. Breaking Ties: Regression Discontinuity Design Meets Market Design. *Econometrica* Forthcoming.

Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda. 2015a. Expanding "choice" in school choice. *American Economic Journal: Microeconomics* 7(1):1–42.

———. 2015b. Expanding "choice" in school choice. *American Economic Journal: Microeconomics* 7(1):1–42.

Abdulkadiroğlu, Atila, Parag A Pathak, and Alvin E Roth. 2005. The new york city high school match. *American Economic Review* 95(2):364–367.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2005. The New York City High School Match. *American Economic Review Papers and Proceedings* 95: 364–367.

———. 2009. Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match. *American Economic Review* 99(5):1954–1978.

Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. 2020. Do parents value school effectiveness? *American Economic Review* 110(5):1502–39.

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak. 2014. The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica* 82(1): 137–196.

Agarwal, Nikhil, and Paulo Somaini. 2018a. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica* 86(2):391–444.

———. 2018b. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica* 86(2):391–444.

———. 2020. Revealed preference analysis of school choice models. *Annual Review of Economics* 12:471–501.

Agostinelli, Francesco, Margaux Luflade, and Paolo Martellini. 2021. On the spatial determinants of educational access. *HCEO Working Paper*.

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. 2015. The economics of density: Evidence from the berlin wall. *Econometrica* 83(6): 2127–2189.

Ajayi, Kehinde, and Modibo Sidibe. 2020. School choice under imperfect information. *Economic Research Initiatives at Duke (ERID) Working Paper* (294).

Ajayi, Kehinde F., Willa H. Friedman, and Adrienne M. Lucas. 2017. The importance of information targeting for school choice. *American Economic Review* 107(5): 638–643.

Allende, Claudia. 2019. Competition under social interactions and the design of education policies. *Working Paper*.

Allende, Claudia, Francisco Gallego, and Christopher Neilson. 2019. Approximating the equilibrium effects of informed school choice. *Work. Pap., Princeton Univ., Princeton, NJ*.

Almagro, Milena, and Tomás Domınguez-Iino. 2019. Location sorting and endogenous amenities: Evidence from amsterdam. *Working Paper*.

Altonji, J.G., P. Arcidiacono, and A. Maurel. 2016. The analysis of field choice in college and graduate school. In *Handbook of the economics of education*, 305–396. Elsevier.

Anderson, Drew M. 2020. When financial aid is scarce: The challenge of allocating college aid where it is needed most. *Journal of Public Economics* 190:104253.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Arcidiacono, Peter, and Paul B. Ellickson. 2011. Practical methods for estimation of dynamic discrete choice models. *Annual Review of Economics* 3(1):363–394.

Arcidiacono, Peter, and John Bailey Jones. 2003. Finite mixture distributions, sequential likelihood and the em algorithm. *Econometrica* 71(3):933–946.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2017. Strategic 'mistakes': Implications for market design research. *Working Paper*.

———. 2021. Strategic 'Mistakes': Implications for Market Design Research. *Working Paper*.

Avery, Christopher, and Parag A Pathak. 2021. The distributional consequences of public school choice. *American Economic Review* 111(1):129–52.

Azevedo, Eduardo M, and Jacob D Leshno. 2016a. A supply and demand framework for two-sided matching markets. *Journal of Political Economy* 124(5):1235–1268.

———. 2016b. A supply and demand framework for two-sided matching markets. *Journal of Political Economy* 124:1235–1268.

Bailey, Martha J, and Andrew Goodman-Bacon. 2015. The war on poverty's experiment in public medicine: Community health centers and the mortality of older americans. *American Economic Review* 105(3):1067–1104.

Barwick, Panle Jia, Shanjun Li, Andrew R Waxman, Jing Wu, and Tianli Xia. 2021. Efficiency and equity impacts of urban transportation policies with equilibrium sorting. Tech. Rep., National Bureau of Economic Research.

Baum-Snow, Nathaniel, and Fernando Ferreira. 2015. Causal inference in urban and regional economics. In *Handbook of regional and urban economics*, vol. 5, 3–68. Elsevier.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy* 115(4):588–638.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins. 2016. A dynamic model of demand for houses and neighborhoods. *Econometrica* 84(3): 893–942.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society* 841–890.

———. 2004. Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy* 112(1): 68–105.

Bettinger, Eric, Oded Gurantz, Laura Kawano, Bruce Sacerdote, and Michael Stevens. 2019. The long-run impacts of financial aid: Evidence from California's Cal Grant. *American Economic Journal: Economic Policy* 11(1):64–94.

Betts, Julian R. 1995. Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth. *The Review of Economics and Statistics* 77(2):231–250.

Bhat, Chandra R. 2000. Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. *Transportation Science* 34(2):228–238.

Billings, Stephen B, Eric J Brunner, and Stephen L Ross. 2018. Gentrification and failing schools: The unintended consequences of school choice under nclb. *Review of Economics and Statistics* 100(1):65–77.

Black, Sandra E. 1999. Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics* 114(2):577–599.

Blagg, Kristin, Matthew Chingos, and Victoria Lee. 2017. The price of room and board: Understanding trends in on-campus living charges.

Blom, Erica, Kelia Washington, Macy Rainer, Carina Chien, and Robert Kelchen. 2020. Ipeds finance user guide.

Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner. 2019. Public universities: The supply side of building a skilled workforce. *RSF: The Russell Sage Foundation Journal of the Social Sciences* 5(5):43–66.

———. 2020. A passage to america: University funding and international students. *American Economic Journal: Economic Policy* 12(1):97–126.

Bound, John, and Sarah Turner. 2007. Cohort crowding: How resources affect collegiate attainment. *Journal of Public Economics* 91(5-6):877–899.

Brunner, Eric J, Sung-Woo Cho, and Randall Reback. 2012. Mobility, housing markets, and schools: Estimating the effects of inter-district choice programs. *Journal of Public Economics* 96(7-8):604–614.

Burdick-Will, Julia. 2017. Neighbors but not classmates: Neighborhood disadvantage, local violent crime, and the heterogeneity of educational experiences in chicago. *American Journal of Education* 124(1):37–65.

Callaway, Brantly, and Pedro HC Sant'Anna. 2020. Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Calsamiglia, Caterina, Chao Fu, and Maia Güell. 2020. Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy* 128(2):642–680.

Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn. 2010. Constrained school choice: An experimental study. *American Economic Review* 100(4):1860–1874.

Card, D., and A. B. Krueger. 1992a. School quality and black-white relative earnings: A direct assessment. *The Quarterly Journal of Economics* 107(1):151–200.

Card, David. 1993. Using geographic variation in college proximity to estimate the return to schooling. *National Bureau of Economic Research Working Paper*.

Card, David, and Alan B. Krueger. 1992b. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy* 100(1):1–40.

Che, Yeon-Koo, Dong Woo Hahm, and YingHua He. 2021. Leveraging uncertainties to infer preferences: Robust analysis of school choice. *Working Paper*.

Che, Yeon-Koo, Dong Woo Hahm, and Yinghua He. 2022. Leveraging uncertainties to infer preferences: Robust analysis of school choice. *Working Paper*.

Che, Yeon-Koo, and Olivier Tercieux. 2019a. Efficiency and stability in large matching markets. *Journal of Political Economy* 127(5):2301–2342.

———. 2019b. Efficiency and stability in large matching markets. *Journal of Political Economy* 127(5):2301–2342.

Chen, Yan, and YingHua He. 2021a. Information acquisition and provision in school choice: a theoretical investigation. *Economic Theory*.

Chen, Yan, and Yinghua He. 2021b. Information acquisition and provision in school choice: An experimental study. *Journal of Economic Theory* 197:105345.

———. 2021c. Information acquisition and provision in school choice: an experimental study. *Journal of Economic Theory* Forthcoming.

Christensen, Peter, Ignacio Sarmiento-Barbieri, and Christopher Timmins. 2020. Housing discrimination and the toxics exposure gap in the united states: Evidence from the rental market. *The Review of Economics and Statistics* 1–37.

Christensen, Peter, and Christopher Timmins. 2018. Sorting or steering: Experimental evidence on the economic effects of housing discrimination. Tech. Rep., National Bureau of Economic Research.

Clark, Damon, and Emilia Del Bono. 2016. The long-run effects of attending an elite school: Evidence from the united kingdom. *American Economic Journal: Applied Economics* 8(1):150–176.

Clelan, Elizabeth B, and Michael S Kofoed. 2017. The Effect of the Business Cycle on Freshman Financial Aid. *Contemporary Economic Policy* 35(2):253–268.

Cohen, Danielle. 2021. Nyc school segregation report card: Still last, action needed now! *UCLA Civil Rights Project*.

Corcoran, Sean, and Henry M Levin. 2011a. School choice and competition in the new york city schools. In *Education reform in new york city: An ambitious change in the nation's most complex school system*. Harvard University Press.

Corcoran, Sean P. 2018. School choice and commuting.

Corcoran, Sean P., Jennifer Jennings, Sarah Cohodes, and Carolyn Sattin-Bajaj. 2018a. Leveling the playing field for high school choice: Results from a field experiment of informational interventions. Tech. Rep.

Corcoran, Sean P, Jennifer L Jennings, Sarah R Cohodes, and Carolyn Sattin-Bajaj. 2018b. Leveling the playing field for high school choice: Results from a field experiment of informational interventions. Tech. Rep., National Bureau of Economic Research.

Corcoran, Sean P., and Henry M. Levin. 2011b. School Choice and Competition in the New York City Schools. In *Education Reform in New York City: Ambitious Change* *in the Nation's Most Complex School System*, ed. Jennifer O'Day, Catherine Bitter, and Louis M. Gomez. Harvard Education Press.

Cunha, Flavio, and James Heckman. 2007. The technology of skill formation. *American Economic Review* 97(2):31–47.

Davis, Morris A, Jesse Gregory, and Daniel A Hartley. 2019. The long-run effects of low-income housing on neighborhood composition.

Davis, Morris A, Jesse Gregory, Daniel A Hartley, and Kegon TK Tan. 2021. Neighborhood effects and housing vouchers. *Quantitative Economics* 12(4):1307–1346.

De Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020a. Difference-in-differences estimators of intertemporal treatment effects. *Available at SSRN 3731856*.

———. 2020b. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110(9):2964–96.

Dehejia, Rajeev H, and Sadek Wahba. 2002. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics* 84(1):151–161.

Delaney, Jennifer A., and Tyler D. Kearney. 2015. The impact of guaranteed tuition policies on postsecondary tuition levels: A difference-in-difference approach. *Economics of Education Review* 47:80–99.

———. 2016. Alternative student-based revenue streams for higher education institutions: A difference-in-difference analysis using guaranteed tuition policies. *Journal of Higher Education* 87:731–769.

Deming, David, and Chris Walters. 2017. The impacts of price and spending subsidies on us postsecondary attainment. *NBER Working Paper* 23736.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2014. School Choice, School Quality, and Postsecondary Attainment. *American Economic Review* 104(3):991–1013.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39(1):1–22.

Dinerstein, Michael, and Troy D Smith. 2021. Quantifying the supply response of private schools to public policies. *American Economic Review* 111(10):3376–3417.

Dinerstein, Michael F, Caroline M Hoxby, Jonathan Meer, and Pablo Villanueva. 2014. Did the Fiscal Stimulus Work for Universities? In *How the financial crisis and great recession affected higher education*, ed. Jeffrey R. Brow and Caroline M. Hoxby, 263–320. University of Chicago Press.

Dingel, Jonathan I, and Felix Tintelnot. 2020. Spatial economics for granular settings. Tech. Rep., National Bureau of Economic Research.

Dobbie, Will, and Roland G. Fryer. 2014. The impact of attending a school with high-achieving peers: Evidence from the new york city exam schools. *American Economic Journal: Applied Economics* 6(3):58–75.

Dur, Umut M, Scott Duke Kominers, Parag A Pathak, and Tayfun Sönmez. 2013. The demise of walk zones in boston: Priorities vs. precedence in school choice. Tech. Rep., National Bureau of Economic Research.

Dynarski, Susan, and Judith Scott-Clayton. 2013. Financial aid policy: Lessons from research. *The Future of Children* 23(1):67–91.

Dynarski, Susan M. 2003. Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review* 93(1):279–288.

Ellen, Ingrid Gould, Keren Mertens Horn, and Amy Ellen Schwartz. 2016. Why don't housing choice voucher recipients live near better schools? insights from big data. *Journal of Policy Analysis and Management* 35(4):884–905.

Ellison, Glenn, and Parag A. Pathak. 2021. The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from chicago's exam schools. *American Economic Review* 111(3):943–975.

Epple, Dennis, Richard Romano, and Holger Sieg. 2006. Admission, tuition, and financial aid policies in the market for higher education. *Econometrica* 74(4): 885–928.

Epple, Dennis N, and Richard Romano. 2003. Neighborhood schools, choice, and the distribution of educational benefits. In *The economics of school choice*, 227–286. University of Chicago Press.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2019a. Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review* 109(4):1486–1529.

———. 2019b. Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions. *American Economics Review* 109(4):1486–1529.

Ferreira, Fernando V, and Maisy Wong. 2020. Estimating preferences for neighborhood amenities under imperfect information. Tech. Rep., National Bureau of Economic Research.

Ferreyra, Maria Marta. 2007. Estimating the effects of private school vouchers in multi district economies. *American Economic Review* 97(3):789–817.

Fethke, Gary. 2005. Strategic determination of higher education subsidies and tuitions. *Economics of Education Review* 24(5):601–609.

Fu, Chao. 2014. Equilibrium tuition, applications, admissions, and enrollment in the college market. *Journal of Political Economy* 122(2):225–281.

Fu, Chao, Shoya Ishimaru, and John Kennan. 2019. Government expenditure on the public education system. *NBER Working Paper* 26425.

Gale, David, and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69(1):9–15.

Grenet, Julien, Yinghua He He, and Dorothea Kübler. 2021. Decentralizing centralized matchingmarkets: Implications from early offers in university admissions. *Working Paper*.

Grigoryan, Aram. 2021. School choice and the housing market. *Available at SSRN 3848180*.

Haeringer, Guillaume, and Flip Klijn. 2009a. Constrained school choice. *Journal of Economic theory* 144(5):1921–1947.

———. 2009b. Constrained school choice. *Journal of Economic Theory* 144(5): 1921–1947.

Hahm, Dong Woo, and Minseon Park. 2022. A dynamic framework of school choice: Effects of middle schools on high school choice. In *Proceedings of the 23rd acm conference on economics and computation*, 292–293.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1):201–209.

Hassidim, Avinatan, Assaf Romm, and Ran I Shorrer. 2016. " strategic" behavior in a strategy-proof environment. In *Proceedings of the 2016 acm conference on economics and computation*, 763–764.

Hastings, Justine S., Thomas Kane, and Douglas Staiger. 2005. Parental preferences and school competition: Evidence from a public school choice program. Tech. Rep.

Hastings, Justine S, and Jeffrey M Weinstein. 2008a. Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics* 123(4):1373–1414.

Hastings, Justine S., and Jeffrey M. Weinstein. 2008b. Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *The Quarterly Journal of Economics* 123(4):1373–1414.

He, Yinghua. 2015. Gaming the boston school choice mechanism in beijing. *Working Paper*.

———. 2017. Gaming the Boston School Choice Mechanism in Beijing.

He, YingHua, Shruti Sinha, and Xiaoting Sun. 2021. Identification and estimation in many-to-one two-sided matching without transfers. *arXiv preprint arXiv:2104.02009*.

Heckman, James, and Burton Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society* 271–320.

Heckman, James J. 2007. The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences* 104(33):13250–13255.

Heiss, Florian, and Viktor Winschel. 2008. Likelihood approximation by numerical integration on sparse grids. *Journal of Econometrics* 144(1):62–80.

Hoekstra, Mark. 2009. The effect of attending the flagship state university on earnings: A discontinuity-based approach. *Review of Economics and Statistics* 91(4):717–724.

Hoxby, Caroline M., and Jonah E. Rockoff. 2004. The impact of charter schools on student achievement. *Working Paper*.

Hoxby, Caroline M, and Sarah Turner. 2015a. What high-achieving low-income students know about college. *American Economic Review* 105(5):514–17.

Hoxby, Caroline M., and Sarah Turner. 2015b. What high-achieving low-income students know about college. *American Economic Review* 105(5):514–517.

Iceland, John. 2004. The multigroup entropy index (also known as theil's h or the information theory index). *US Census Bureau. Retrieved July* 31:2006.

Idoux, Clemence. 2022. Integrating new york city schools: The role of admission criteria and family preferences. Tech. Rep., Technical Report.

Imbens, Guido, and Karthik Kalyanaraman. 2012. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies* 79(3): 933–959.

Jackson, Kirabo C. 2010. Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. *The Economic Journal* 120(549):1399–1429.

Jann, Ben. 2017. Kmatch: Stata module for multivariate-distance and propensity-score matching.

Jaquette, Ozan, and Edna Parra. 2016. The problem with the delta cost project database. *Research in Higher Education* 57(5):630–651.

Kelchen, Robert. 2016. An analysis of student fees: The roles of states and institutions. *The Review of Higher Education* 39(4):597–619.

Klarner, Carl. 2018. State legislative election returns, 1967-2016. *https://doi.org/10.7910/DVN/3WZFK9, Harvard Dataverse, V3, UNF:6:pV4h1CP/B8pHthjjQThTTw==*.

Kulka, Amrita. 2019. Sorting into neighborhoods: The role of minimum lot sizes. *Working Paper*.

Laverde, Mariana. 2020. Unequal assignments to public schools and the limits of school choice. *Working Paper*.

Li, Shengwu. 2017. Obviously strategy-proof mechanisms. *American Economic Review* 107(11):3257–3287.

Luflade, Margaux. 2017. The value of information in centralized school choice systems. *Working Paper*.

———. 2018. The value of information in centralized school choice systems. *Working Paper*.

Manski, Charles F. 1992. Educational choice (vouchers) and social mobility. *Economics of Education Review* 11(4):351–369.

Mark, Nicholas D.E., Sean P. Corcoran, and Jennifer Jennings. 2021. Choosing alone? peer similarity in high school choices. *Working Paper* (396).

Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J Redding. 2022. The economics of spatial mobility: Theory and evidence using smartphone data. Tech. Rep.

Monarrez, Tomás E. 2020. School attendance boundaries and the segregation of public schools in the u.s. *American Economic Journal: Applied Economics*.

Mountjoy, Jack. 2022. Community colleges and upward mobility. *American Economic Review* 112(8):2580–2630.

Nechyba, Thomas J. 2000. Mobility, targeting, and private-school vouchers. *American Economic Review* 90(1):130–146.

Parry, Ian WH, and Kenneth A Small. 2009. Should urban transit subsidies be reduced? *American Economic Review* 99(3):700–724.

Pop-Eleches, Cristian, and Miguel Urquiola. 2013. Going to a Better School: Effects and Behavioral Responses. *American Economic Review* 103(4):1289–1324.

Purdue University. 2020. Purdue announces 9th consecutive tuition freeze.

Purevjav, Avralt-Od. 2022. The welfare effects of congestion-reduction policies. *Working Paper*.

Reback, Randall. 2005. House prices and the provision of local public services: Capitalization under school choice programs. *Journal of Urban Economics* 2(57): 275–301.

Rizzo, Michael J., and Ronald G Ehrenberg. 2004. Resident and Nonresident Tuition and Enrollment at Flagship State Universities. In *College choices: The economics of where to go, when to go, and how to pay for it*, 303–354. University of Chicago Press.

Rosenbaum, Paul R, and Donald B Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55.

Rust, John. 1994. Structural estimation of markov decision processes. *Handbook of Econometrics* 4:3081–3143.

Sandler, Danielle H., and Ryan Sandler. 2014. Multiple event studies in public finance and labor economics: A simulation study with applications. *Journal of Economic and Social Measurement* 39(1-2):31–57.

Sattin-Bajaj, Carolyn, and Jennifer Jennings. 2022. Understanding safety as a driver of school selection for parents and students: Lessons from new york city. In *2021 appam fall research conference*. APPAM.

Sattin-Bajaj, Carolyn, Jennifer L Jennings, Sean P Corcoran, Elizabeth Christine Baker-Smith, and Chantal Hailey. 2018. Surviving at the street level: How counselors' implementation of school choice policy shapes students' high school destinations. *Sociology of Education* 91(1):46–71.

Schwartz, Amy Ellen, Leanna Stiefel, and Matthew Wiswall. 2013. Do small schools improve performance in large, urban districts? causal evidence from new york city. *Journal of Urban Economics* 77:27–40.

Schwartz, Amy Ellen, Ioan Voicu, and Keren Mertens Horn. 2014. Do choice schools break the link between public schools and property values? evidence from house prices in new york city. *Regional Science and Urban Economics* 49:1–10.

Singell Jr, Larry D, and Joe A Stone. 2007. For whom the pell tolls: The response of university tuition to federal grants-in-aid. *Economics of Education review* 26(3): 285–295.

Smith, Jeffrey A, and Petra E Todd. 2005. Does matching overcome lalonde's critique of nonexperimental estimators? *Journal of econometrics* 125(1-2):305–353.

Son, Suk Joon. 2020. Distributional impacts of centralized school choice. *Working Paper*.

Sun, Liyang, and Sarah Abraham. 2020. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

Train, Kenneth E. 2009. *Discrete choice methods with simulation*. Cambridge university press.

Turner, Lesley J. 2017. The economic incidence of federal student grant aid. *Working Paper*.

U.S. Bureau of Labor Statistics. 2021. Local area unemployment statistics.

Varadhan, Ravi, and Christophe Roland. 2008. Simple and globally convergent methods for accelerating the convergence of any em algorithm. *Scandinavian Journal of Statistics* 35(2):335–353.

Walters, Christopher R. 2018. The demand for effective charter schools. *Journal of Political Economy* 126(6):2179–2223.

Webber, Douglas A. 2017. State divestment and tuition at public institutions. *Economics of Education Review* 60:1–4.

Xu, Jing. 2019. Housing choices, sorting, and the distribution of educational benefits under deferred acceptance. *Journal of Public Economic Theory* 21(3):558–595.

Yang, Rebecca, and Paul A Jargowsky. 2006. Suburban development and economic segregation in the 1990s. *Journal of Urban Affairs* 28(3):253–273.

Zheng, Angela. 2022. The valuation of local school quality under school choice. *American Economic Journal: Economic Policy* 14(2):509–37.