Match or Mismatch: Learning and Inertia in School Choice

Yusuke Narita*†

January 20, 2016
Click Here for the Latest Version

Abstract

Centralized matching markets are designed assuming that participants make well-informed choices upfront. However, this paper uses data from NYC's school choice system to show that families' choices change after the initial match as they learn about schools. I develop an empirical model of evolving demand for schools under learning, switching costs, and demand responses to prior assignments. These model components are identified by using admissions lotteries and other institutional features. The estimates suggest that there are even more changes in underlying demand than in observed choices, undermining the welfare performance of the initial match. To alleviate the welfare cost of demand changes, I theoretically and empirically investigate dynamic mechanisms that best accommodate choice changes. These mechanisms improve on the existing discretionary reapplication process. In addition, the gains from the mechanisms change substantially depending on the extent of demand-side inertia caused by switching costs. Thus, the gains from a centralized market depend not only on its design but also on demand-side frictions (such as demand changes and inertia).

*Massachusetts Institute of Technology. narita@mit.edu.

†I am most grateful to Parag Pathak for continuous encouragement, help, and suggestions from the start of this project. Daron Acemoglu, Nikhil Agarwal, and Josh Angrist provided numerous conceptual and technical comments as well as criticisms, which were crucial for the current form of this paper. For helpful discussions, I also thank David Autor, Abhijit Banerjee, Esther Duflo, Glenn Ellison, Jon Gruber, Jerry Hausman, James Poterba, many MIT and Harvard students, especially Alex Bartik, Vivek Bhattacharya, Alonso Bucarey, Masao Fukui, Chishio Furukawa, Ludo Gazze, Mark Hou, Sally Hudson, Peter Hull, Tetsuya Kaji, Yuhei Miyauchi, Manisha Padi, Brendan Price, Elizabeth Setren, Kentaro Tomoeda, and Yufei Wu. Special thanks go to Louis Chen and Daniel Waldinger for careful proof-reading. Financial support from the Shultz fund at MIT is gratefully acknowledged.

1 Introduction

From public housing to entry-level labor markets to school choice, centralized matching markets are a prominent form of public policy. These markets are usually designed assuming that participants make well-informed choices upfront, anticipating what they will demand in the future. Little is known, however, about whether this assumption holds in practice. In this paper, I study how families' demand for schools evolves using data from NYC's high school choice system, which serves more than 90,000 families applying for 700 schools each year. I develop a framework to recover families' evolving demand under learning and switching costs, where demand is also allowed to change in response to prior assignments. I use the estimated framework to quantify demand changes and their welfare consequences.

My analysis uses administrative panel data from NYC's high school choice system, consisting of the centralized first-round in December and a discretionary reapplication process in April. The data records families' participation decisions as well as rank-ordered school choices both in the first-round and the reapplication process, allowing me to trace the dynamics of school choices. In Section 2, I demonstrate that contrary to the premise of well-informed upfront choices, families' choices change after the initial match as they learn about schools. About 7% of families reapply and at least 70% of these reapplicants reverse choice (preference) orders over schools between the first round and the reapplication process. Most choice reversals can be rationalized only by real demand changes and not by strategic behavior with unchanged demand.[1] These choice changes appear to be mainly caused by learning. Most families self-report that they change their choices because of new information about schools or their preferences about schools. Moreover, consistent with their self-reports, families' choices become more correlated with and responsive to school characteristics. In particular, compared with initial applications, reapplications rank schools that are closer to families' homes and that are academically no worse.

The above choice changes provide only a lower bound on the amount of changes in underlying demand. There may or may not be additional *un*observed demand changes. For families who do not reapply, their behavior does not directly reveal demand changes. However, some families may experience demand changes but not reapply because of "switching

[1] I say an applicant exhibits choice reversals if the following holds: In the reapplication process, she attempts to switch from the first-round assignment s to another school that is unranked or ranked below s in the preference list she reports in the first-round market. I say an applicant exhibits surely nonstrategic choice reversals if she exhibits choice reversals and does not exhaust her first-round preference, i.e., rank 11 or fewer schools though she could have ranked up to 12 schools. I show that (1) 71% of reapplicants exhibit choice reversals, (2) about 80% of the choice reversals are surely nonstrategic, and (3) surely nonstrategic choice reversals are consistent with optimal behavior only if intrinsic preferences and demand change between the first round and the reapplication process.

costs" or "reapplication costs," in which case demand changes exist but are not directly observed. For example, the time cost of filling out a paper application form may constitute reapplication costs.

To distinguish these scenarios and recover underlying evolving demand, the core part of this paper in Section 3 develops a structural model of dynamic school choice. The model incorporates demand changes by learning and reapplication costs. I also let demand change in response to initial assignments in the first round. Specifically, I allow the utility of the first-round assignment to change by some positive or negative amount, which can differ across applicants. Such demand responses are likely if families obtain more information about their initial assignments or if they experience psychological endowment effects about them. I provide a strategy to distinguish these behavioral elements by exploiting three institutional features. First, because of capacity constraints, many applicants are initially assigned to a school other than the most preferred school. Second, the assignments to the most preferred and less preferred schools are partly random, thanks to admissions lotteries used in the first-round assignment mechanism. Finally, reapplicants make new choices that are rank-ordered. These features allow me to decompose observed behavior into the model components (demand changes, reapplication costs, and demand changes in response to prior assignments.).

Intuitively, the identification logic is as follows. For simplicity, let me ignore demand responses to initial assignments and assume that there are only two preference ranks, the first choice and the lower choice. Many applicants are "lower-choice non-reapplicants," who are initially assigned to their old lower choice but do not reapply, due either to reapplication costs or demand changes. This fact allows me to measure the total effects of demand changes and reapplication costs. There are also "first-choice reapplicants," who are initially assigned to their old first choice but reapply, which must be because of demand changes. Other applicants assigned to their first choice may also experience demand changes but be locked in by reapplication costs. The fraction of reapplicants among all applicants assigned to the first choice thus tells us the difference between the amounts of demand changes and reapplication costs.

Now suppose that admissions lotteries in the first-round mechanism guarantee that initial assignments are randomly assigned and applicants assigned to the first-choice and lower-choice are comparable people with similar demand changes and reapplication costs. I can compare the fractions of lower-choice non-reapplicants and first-choice reapplicants to measure the amount of reapplication costs. Heuristically,

(fraction of non-reapplicants among applicants assigned to the initial lower choice)

-(fraction of reapplicants among applicants assigned to the initial first choice)
=(demand changes+reapplication costs)-(demand changes-reapplication costs)
=2×reapplication costs,

which separates reapplication costs from demand changes. The precise implementation of this argument generates complications related to more than two preference ranks, admissions lotteries embedded in assignment mechanisms, and demand responses to initial assignments. Solutions for these challenges are detailed below.

I estimate the model and find a significant role of learning, reapplication costs, and demand responses to initial assignments. Crucially, as detailed in Section 4, the estimates suggest that there are substantially more changes in underlying demand than in observed choices. These hidden demand changes are masked by reapplication costs, which prevent families from reapplying and expressing demand changes.[2]

As a result, the welfare cost of ignoring demand changes is large. To measure the welfare cost, I compare the real first-round assignment based on old demand with the counterfactual "frictionless benchmark." The frictionless benchmark is defined as what would have been produced by the same first-round assignment mechanism, had families made choices based on their new demand after learning. Since the two differ only in whether families' choices are based on old or new demand, the difference between the two captures the welfare costs of ignoring demand changes by learning.[3] The real and frictionless assignments turn out to be significantly different. Specifically, the two assignments give different allocations (schools) to a majority of families; the average welfare loss under the real first-round assignment compared with the frictionless benchmark is more than 1-mile-equivalent, when I measure it by new demand assumed to be quasi-linear in the distance between the family and the school locations.[4] This magnitude corresponds to more than .15 standard deviations in the distribution of utilities from schools for each applicant. Demand changes thus undermine the welfare performance of the initial match that ignores demand changes.

The large welfare cost of ignoring demand changes motivates me to investigate ways to alleviate the cost by accommodating demand changes. As already explained, NYC runs a

[2]Demand responses to initial assignments also lower the reapplication rate. The estimates show that families tend to get to prefer initially assigned schools more, compared with other schools. This satisfaction with initial assignments lowers the reapplication rate.

[3]Except applicants' choices, every other input is unchanged between the real first-round assignment and the frictionless benchmark. For example, school capacities and their preferences or priorities over applicants are fixed.

[4]The 1-mile-equivalent utility unit can be interpreted as corresponding to traveling 1 mile every school day during the high school years. Also, I use new demand as my welfare measure because it is demand after leaning at a point in time closer to enrollment periods; new demand is thus expected to be a better welfare measure than old demand is.

discretionary, human-driven reapplication process. It is also possible to run a centralized algorithm for the reapplication process. I build a dynamic version of the school-student assignment model, analyze centralized designs of the reapplication process, and show that the centralized reapplication processes are the "best possible" mechanisms to accommodate choice changes. I evaluate how well the discretionary and centralized reapplication processes alleviate the welfare cost of ignoring demand changes. I find that both types of reapplication processes produce welfare gains, but the centralized reapplication processes are more effective and produce gains more than twice as large as those from the discretionary process.

This evaluation of reapplication processes takes estimated reapplication costs as given. There are technological changes and school districts' and social entrepreneurs' initiatives that may ease reapplication costs (e.g., online systems for more easily making and updating school choices). To measure the potential effects of such demand-side changes or interventions, I finally investigate how the performance of reapplication processes depends on reapplication costs and the resulting demand-side inertia. I find that the gains from the centralized reapplication mechanisms change by several times depending on the extent of demand-side inertia, which governs how much demand changes are revealed in reapplications.

These findings show that learning causes significant demand changes, which in turn undermine the welfare performance of the initial match and result in the large welfare cost of ignoring demand changes. Dynamic reapplications processes, especially centralized ones, help alleviate the welfare loss by accommodating changing demand. In addition, the gains from the mechanisms substantially change depending on the extent of demand-side inertia caused by reapplication costs. Thus, in the dynamic real world, the gains from a centralized market depend not only on its design but also on demand-side frictions such as demand changes (arising from learning) and inertia (caused by reapplication costs). This sheds empirical light on the potential importance of demand-side interventions that attempt to alleviate these frictions (e.g., applications for more easily searching school characteristics, online systems for more easily updating school choices).[5]

Related Literature. This paper is a first empirical study on the welfare performance of a dynamic centralized matching market. For that purpose, I develop an empirical model of evolving demand for schools under learning, reapplication costs, and demand changes in

[5]Many school districts and social entrepreneurs have been launching such initiatives. For example, see `http://www.dnainfo.com/new-york/20131113/washington-heights/six-apps-launch-guide-families-through-high-school-admissions-process` and `http://izonenyc.org/initiatives/innovate-nyc-schools/#scdc`. However, aside from a few field experiments in Hastings and Weinstein (2008) and Andrabi et al. (2015), little is known about the effects of such demand-side interventions.

response to prior assignments, which have been studied in labor and public economics and industrial organization. I propose a novel approach to identify and estimate the different model elements using institutional features of centralized school choice systems. Finally, I combine the estimated model with a theoretical analysis of dynamic centralized matching markets to conduct welfare analysis.

More specifically, this paper combines and contributes to four different strands of the literature. First, the descriptive analysis of evolving school choices uses a revealed preference idea and relates to the economics and psychology literatures on revealed preference changes in the field, e.g., see papers reviewed in Blundell (1988) and Varian (2006) as well as more recent papers including Abaluck and Gruber (2011), Echenique et al. (2011), and Choi et al. (2014). These papers mainly focus on econometric difficulties in the detection and interpretation of preference changes.[6]

Second, my analysis suggests that school choice changes appear to be primarily caused by frictions in initial choices and learning about schools. This suggested importance of frictions in school choices echoes existing studies such as Hastings and Weinstein (2008), Jochim et al. (2014), Andrabi et al. (2015), Wiswall and Zafar (2015), and Hastings et al. (2015).[7] Similar findings are also present in non-education contexts (Fang et al., 2008; Kling et al., 2012; Handel and Kolstad, 2015). Unlike these studies, I study the dynamics of frictions and learning within the same applicant or family. In this respect, this paper relates to non-education papers like Farber and Gibbons (1996), Ketcham et al. (2012), and Ketcham et al. (2015) on learning.

Third, my empirical model is a model of dynamic school choice with frictions in initial choices, reapplication costs, and demand responses to initial assignments. My model is thus at the intersection of empirical school choice models, demand models with frictions, and those with switching costs (or "state dependence" or "brand loyalty"), which have heretofore been applied only to non-education settings such as labor supply, insurance, and

[6] I am more interested in welfare consequences of demand changes in dynamic centralized markets. To my knowledge, few papers investigate any pervasive effects of changes during administrative time delays. An exception is Autor et al. (2015), who study the effect of administrative decision time on the labor force participation and earnings of disability insurance applicants. Also, existing studies on *de*centralized matching markets emphasize the importance of demand changes in their discussion of unravelling. See Roth (2002) for an overview.

[7] NYC's former deputy director of high school enrollment also points to potential frictions in the school choice process: "Given how massive the New York City process is, (...) the process by which those choices are made remains complicated, and very much depends on expertise or the ability to spend an excessive amount of time understanding how it works. Many students still go without either." (`http://ny.chalkbeat.org/2015/08/07/why-high-school-admissions-actually-doesnt-work-for-many-city-students-and-how-it-could/#.VdP8ZHj5q20`)

retail industrial organization.[8] Existing models of switching costs, state dependence, and brand loyalty include Shum (2004), Card and Hyslop (2005), papers reviewed in Farrell and Klemperer (2007), Dube et al. (2009), Bronnenberg et al. (2012), Abaluck and Gruber (2013), Handel (2013), Ho et al. (2015), Sudhir and Yang (2014), and Hortacsu et al. (2015). Demand under choice frictions is studied by Chetty (2012), Handel and Kolstad (2015), and Spinnewijn (2015) among others. Models of school choice include Manski and Wise (1983), Epple et al. (2006), Bayer et al. (2007), Hastings et al. (2008), Ajayi (2013), Neilson (2013), Casalmiglia et al. (2014), He (2014), Walters (2014), Abdulkadiroğlu et al. (2015), Agarwal and Somaini (2015), Dinerstein and Smith (2015), and Kapor (2015).

Finally, this paper's theoretical analysis of centralized dynamic reapplication processes relates to the literature on the design of school-student assignment mechanisms, initiated by Abdulkadiroğlu and Sönmez (2003), building on the classic two-sided matching problem (Roth and Sotomayor, 1990). Their model is extended and applied by Abdulkadiroğlu et al. (2009) to analyze the NYC institution, but both papers consider only static models and avoid dynamic considerations. By contrast, I use a dynamic model to take demand changes into considerations. In this sense, this paper's theoretical analysis shares some attributes with recent papers on dynamic aspects of matching market design, e.g., Ünver (2010), Pereyra (2013), Dur and Kesten (2014), Kennes et al. (2014), Kurino (2014), Anderson et al. (2015), Akbarpour et al. (2015), Baccara et al. (2015), Leshno (2015), and Kadam and Kotowski (2015). This paper differs, however, because these studies exclude demand changes or use additional structures (e.g., binary preferences) to analyze their applications, which makes it difficult to apply them to analyze this paper's problem. More importantly, none of the above theoretical papers connects theory to data.

2 A First Look at Evolving Choices

2.1 Evolving School Choices in NYC

I start by documenting how families' school choices evolve over time. My analysis uses administrative panel data from the public high school choice system in NYC for the 2004-5 school year. This system contains more than 700 high school programs of various types across Greater New York. Some schools are academically selective, while others put emphasis on

[8] My model is also related to demand models with learning (see Ching et al. 2013 for a review), but different in that these learning models usually consider forward-looking consumers who try to learn the quality of frequently-used products (e.g., detergents) by experimenting with multiple products. This learning-by-experimentation aspect seems secondary in my context of education, where it is not easy to switch from one school to another.

the arts. 8th (and some 9th) graders living in NYC may apply to these schools. Each year, about 90,000 families apply, and most of them are admitted by some school. This system has been organizing applications and selections via the following centralized procedure:

(1) Each applicant ranks up to 12 schools in the order of her preference.

(2) Each school ranks applicants using its preference or priority as well as lottery numbers.

(3) NYC runs a strategy-proof algorithm (the "deferred acceptance" algorithm) on applicants' and schools' preferences to make an initial assignment of applicants to schools.[9]

On top of this initial match, the system has an additional reapplication process. After being informed of the initially assigned school, any applicant is allowed to reapply against the assigned school if she is not satisfied with it. A reapplicant needs to fill out a paper reapplication form with a written reason for reapplying and turn it in to the guidance counselor at her middle school. In the reapplication, she is asked to rank up to 3 other schools she currently prefers over the initial assignment. She can rank the same schools as in the initial application. The initial assignment is guaranteed, i.e., if her reapplication is rejected, she is assigned her initially assigned school.

The timeline of initial applications and reapplications is available in Figure 1. Applicants make initial applications during November and December. After the announcement of the initial assignment, some families file reapplications during April and May. For reapplicants, the time interval between their initial application dates and reapplication dates are of mean 153.8 days and standard deviation 7.7 days (Appendix Figure A.1). Thanks to the reapplication process, for those who reapply, I observe their school choices at two different points in time, which allows me to investigate how their demand for schools evolves.[10]

About 7% (6430 applicants) of 91289 applicants reapply (Table 1). NYC accepts and reassigns 21% of reapplications to other schools in a discretionary, human-driven reapplication

[9] The details of the algorithm will be explained in Appendix A.4.

[10] After the initial application process described below, there is the "supplementary round" for students who are not matched in the initial match. I exclude the supplementary round from the analysis because the supplementary round lets students rank only schools they do not rank in the initial application process, which makes it impossible to observe any clear choice changes or reversals between the initial application process and the supplementary round. In addition, the separate system that NYC uses for allocating seats in selective "specialized programs" or exam schools, is also outside the scope of my analysis because the system does not provide information about dynamic choice changes. I also exclude from my analysis those applicants who enroll in schools other than their initial assignments through over-the-counter bargaining, because I observe little information about it. Finally, my description below is for school year 2004-5 and parts of it may not be applicable to the current institution. Nevertheless, NYC keeps using similar discretionary reapplication processes even in recent years. See `http://insideschools.org/blog/item/1000804-kids-win-one-third-of-hs-appeals#`.

process. To measure choice changes, I say an applicant *exhibits choice reversals* if she reapplies against her initially assigned school s by ranking another school that is ranked below s or unranked in her initial application.[11] Among those who reapply, 71% (4564 applicants) exhibit choice reversals. This number is about 5% of the whole population.[12]

Crucially, most choice reversals can be rationalized only by intrinsic demand or preference changes. I say an applicant *exhibits surely nonstrategic choice reversals* if she exhibits choice reversals and ranks 11 or fewer schools in her initial application, even though she could have ranked up to 12 schools. Surely nonstrategic choice reversals are consistent with optimal behavior only if there are intrinsic demand changes between the first round and the reapplication process. To see this, suppose to the contrary that an applicant exhibits surely nonstrategic choice reversals but does not experience demand changes. Let s be her initial assignment and t be any other school that (1) she ranks in her reapplication, but (2) she ranks below s or does not rank in her initial application. If she prefers s to t, then she would be better off by dropping t from her reapplication. If she prefers t to s, then she would gain by ranking t ahead of s in her initial application: The deferred acceptance algorithm in the initial application process is known to be strategy-proof for applicants and guarantees this property (Abdulkadiroğlu and Sönmez, 2003). Thus, surely nonstrategic choice reversals can be rationalized only by real demand changes. Table 1 shows that about 80% of choice reversals are surely nonstrategic. This fact suggests that most choice reversals in the reapplication process reflect real demand changes rather than strategic behavior.

2.2 Choice Frictions and Learning

Characteristics of all applicants, reapplicants, reapplicants who exhibit choice reversals are in Table 2. Those who reapply (and exhibit choice reversals) look similar to the average applicant, though the former is slightly more likely to be a female 8th grader and have lower test scores. Why do these similar looking applicants reapply? There are many potential reasons, such as mistakes in initial applications, changes in the life situation (e.g., moving), changes in the information about schools, changes in intrinsic tastes and preferences, and peer effects related to which schools siblings, friends, and bullies are assigned.

To understand the relative importance of these factors, Figure 2 provides a breakdown of

[11] Note that I do not include choice reversals among other schools than the initially assigned school s. For example, there are cases where an applicant prefers $t(\neq s)$ to $u(\neq s)$ in the initial application, but prefers u to t in the reapplication. I ignore these cases to make my calculation conservative.

[12] As another look at this fact, Appendix Figure A.2b shows that many applicants reapply after being assigned to their top choices. Also, Appendix Figure A.2a shows that the first choice market shares of schools change from the first round to the reapplication process, where the *first choice market share* of a school is the fraction of applicants who rank it first among all applicants who make a first choice.

reapplication reasons self-reported by reapplicants. Panel (a) shows that the vast majority of reapplicants claim that they reapply because of new information or learning about school characteristics or their preferences about school characteristics. For example, many families claim that they were not aware of how far away the initially assigned school is or how painful it is to travel to the initial assigned school. Only a small fraction (less than 5% each) ascribes their reapplications to other potentially important factors, such as mistakes in initial applications and moving after initial applications. This provides suggestive evidence that the main factor for evolving demand is frictions in the initial choice process and learning about schools.[13]

Panel (b) provides a further breakdown of the largest category of new information. This further breakdown shows that a variety of observable and unobservable school characteristics matter. Nevertheless, only a tiny fraction of reapplicants claim that they reapply because they do or do not want to enter the same school as particular siblings or friends or bullies; as far as reapplication decisions are concerned, peer effects do not seem quantitatively important.

Consistent with their self reports, families' choices become more correlated with and responsive to school characteristics in reapplications. Table 3 documents that, compared with initial applications, reapplications rank schools that are more than 20% closer to families' homes. These distance reductions come without sacrificing academic achievement level, as shown in lower rows. This pattern holds across demographic groups (Appendix Table A.2). In Table 4, to incorporate other horizontal characteristics, I run the following descriptive regression:

$$y_s^t = \boldsymbol{b}^t \mathbf{X}_s + e_s^t,$$

where y_s^t is the first choice market share of school s in round t, which is the first round or the reapplication process, where the *first choice market share* of a school is the fraction of applicants who rank it first among all applicants who make a first choice. $\mathbf{X}_s$ is a vector of observable characteristics of school s, which do not change from the first round to the reapplication process.

Table 4 shows R^2's from the above regression for the first round and the reapplication process. Across various specifications of $\mathbf{X}_s$, R^2 is always higher in the reapplication process; reapplications appear to be more attentive or responsive to observable school characteristics than initial applications do. I checked to ensure that this pattern is robust to many specifi-

[13]Since the reapplication process is discretionary, some of the self-reported reapplication reasons may be contaminated by strategic reporting. However, there seems to be no clear reason to expect that strategic reporting overstates the new information category because other reasons, such as moving and mistakes, sound more legitimate.

cations with different school characteristics and their interactions. The R^2 increase is almost always present across demographic groups defined by baseline test scores and race (Appendix Table A.1).[14] Note that Tables 3 and 4 and Figure 2 use mutually exclusive aspects of the data: While Figure 2 classifies self-reported reasons for reapplications, Tables 3 and 4 correlate families' school choice behavior with observable school characteristics. Tables 3 and 4 thus provide another independent support for the possibility that families become more informed of observable school characteristics or their preferences about them as time goes by.

The above descriptive analysis documents that a significant fraction of families change their school choices mainly because of learning about schools. This analysis has several limitations, however. I cannot extrapolate the suggestive findings on learning (Tables 3 and 4 and Figure 2a) to the whole population since they are based on self-selecting reapplicants. More importantly, the descriptive analysis depends entirely on observed choice changes, but observed choice changes may underestimate changes in latent demand. In particular, for families who do not reapply, their behavior does not directly reveal demand changes. However, some families may experience demand changes but not reapply because of "switching costs" or reapplication costs, in which case demand changes exist but are not directly observed. For example, the time cost of filling out a paper application form may constitute a reapplication cost.

To distinguish these scenarios and recover underlying evolving demand, it is necessary to model how learning and the resulting demand changes do or do not come to the surface as observed choice changes in the presence of potential reapplication costs. I next integrate key pieces of the descriptive analysis into a structural model of dynamic school choice with learning and reapplication costs.

3 Uncovering Evolving Demand

3.1 Dynamic School Choice under Learning and Switching Costs

To recover underlying evolving demand that is not necessarily reflected in observed choice behavior, this section develops a structural model of dynamic school choice. The model incorporates demand changes by learning and reapplication costs. In addition, the model allows demand to change in response to initial assignments in the first round. Specifically, I allow the utility of the first-round assignment to change by a positive or negative amount,

[14] These R^2 increases may be trivial if initial applications shares are more dispersed. However, the standard deviation of market shares changes little between the two periods (0.0028 for initial application shares and 0.0027 for reapplication shares).

which can be heterogenous across applicants. Such demand responses are likely if families experience psychological endowment effects about their initial assignments or get more information about them.

Demand Before and After Learning. The random utility of school s for applicant a in period 0 (the initial application process) is

$$U_{as}^0 = U_s^0 + \Sigma_{k=1}^K \beta_{ak}(1 + f_{ak})X_{ask} + \epsilon_{as}^0, \tag{1}$$

where U_s^0 is a school-specific effect, $\mathbf{X}_{as} \equiv (X_{ask})_{k=1,\ldots,K}$ is a vector of (interactions of) a's and s's observable characteristics (e.g., the distance between a's and s's locations), $\boldsymbol{\beta}_a \equiv (\beta_{ak})_{k=1,\ldots,K}$ is a vector of preference coefficients, and ϵ_{as}^0 is an unobserved utility shock.

$\mathbf{f}_a \equiv (f_{ak})_{k=1,\ldots,K}$ is the only non-standard term and stands for frictions a faces about how to value characteristics $\mathbf{X}_{as}$ in the initial application process. I interpret each $\beta_{ak}(1+f_{ak})X_{ask}$ as a's perceived valuation of X_{ask} in $t = 0$. f_{ak} can be positive or negative and heterogenous across different characteristics. Two interpretations of this specification are possible. The first interpretation is that f_{ak} is frictions about preferences β_{ak} and an applicant may not know her preferences about characteristics X_{ask}, e.g., how painful it is to travel a certain distance. The alternative interpretation is that f_{ak} is frictions about characteristics X_{ask} and an applicant may not know X_{ask}, e.g., the distance to schools. Both interpretations and their combinations are consistent with the descriptive analysis and result in the same welfare implications below. I prefer the first interpretation; see "Discussions on Modeling Decisions" at the end of this section for an additional discussion. The modeling of the friction is motivated by Figure 2a and Tables 3 and 4, which suggest that applicants face frictions about observable school characteristics or their preferences about school characteristics in the initial application process. I assume that each applicant a's initial preference $\succ_a^0$ is based on perceived utilities U_{as}^0's subject to frictions, i.e., $s \succ_a^0 s'$ only if $U_{as}^0 > U_{as'}^0$.[15]

After the initial application, NYC runs the (applicant-proposing) deferred acceptance algorithm to give an initially assigned school s_a^0 to each applicant a. During and after the match-making process, applicants' perceived utilities change. The random utility of school s for applicant a in period 1 (the reapplication process) is

$$\bar{U}_{as}^1 = \underbrace{U_s^1 + \boldsymbol{\beta}_a \mathbf{X}_{as} + \epsilon_{as}^1}_{\equiv U_{as}^1} + \gamma_a \mathbb{1}\{s = s_a^0\}. \tag{2}$$

The first three terms are similar to those in initial utilities U_{as}^0 except that the frictions

[15] I assume $s \succ_a^0 s'$ for any ranked school s and unranked school s'. See "Discussions on Modeling Decisions" at the end of this section for discussions about this truth-telling assumption.

are normalized to zero,[16] and school-specific effects $U_s^1 \equiv U_s^0 + U_s$ and unobserved utility shocks $\epsilon_{as}^1 \equiv \epsilon_{as}^0 + \epsilon_{as}$ are subject to new unobserved shocks U_s and ϵ_{as}, respectively. I allow U_s^1 and ϵ_{as}^1 to differ from U_s^0 and ϵ_{as}^0, respectively, to accommodate the fact that demand changes are sometimes related to *un*observed school characteristics such as how nice current students are (Figure 2b). As a result of this specification, unobserved utility shocks ϵ_{as}^0 and ϵ_{as}^1 are serially correlated for each applicant a. This is reasonable given the interpretation of unobserved utility shocks ϵ_{as}^t as the sum of unobserved utility components in period t and that the unobserved determinants of utilities for an applicant are likely to be serially correlated.

The last term $\gamma_a \mathbb{1}\{s = s_a^0\}$, which is turned on if and only if school s is applicant a's initial assignment s_a^0, captures the possibility that an applicant's utility from the initially assigned school may evolve differently than utilities from other schools do. For example, applicants may get more information about the initially assigned school than they would with other schools. Or they may begin to prefer the assigned school more because it admits them, or they get used to it (habit formation or endowment effects). I call γ_a the *initial assignment effect*.

Model of Reapplications 1: Rational Expectation. Each applicant decides whether to reapply based on how preferable the initial assignment s_a^0 is with respect to new demand $\bar{U}_{as}^1$'s. Recall that there are factors that may prevent applicants from reapplying, for example, the time cost of making and submitting a reapplication. In fact, each reapplicant needs to fill out a paper reapplication form with a written reason for reapplying, and turn it in to the guidance counselor at her middle school. I consider two models to incorporate such "reapplication costs" or "switching costs". I call these models the *rational expectation model* and the *naive free expectation model*. The rational expectation model allows for school-specific reapplication acceptance probabilities, but assumes "rational" expectation about applicants' expectations about reapplication acceptance probabilities. The naive free expectation model does not need the rational expectation assumption, but assumes that reapplicants form simplistic beliefs about how the reapplication process works.

The first rational expectation model consists of two layers. The first layer is about reapplication acceptance probabilities. In the reapplication process, each applicant a who reapplies with new preference (s_1, s_2, s_3) is re-assigned to at most one of schools s_1, s_2, and s_3. Since there is no accurate algorithmic description of discretionary reapplication acceptance decisions by NYC, I suppose that the (mutually exclusive) re-assignment probabilities can

[16] In reality, some of the frictions are likely to remain even in the reapplication process. I need to assume it away as a normalization, however, since the data contains only two periods.

be approximated by the following descriptive model. For each $i = 1, 2, 3$,

$$\begin{aligned}
&\Pr(a \text{ is re-assigned to } s_i) \\
&= \begin{cases} \Pr(b_0 + \boldsymbol{b}_X \boldsymbol{X}_{as_i} + \boldsymbol{b}_W \boldsymbol{W}_{as_i} + \xi_{as_i} \geq 0) & \text{if } s_i \neq \emptyset \\ 0 & \text{if } s_i = \emptyset, \end{cases}
\end{aligned}$$

where $\boldsymbol{X}_{as}$ is the characteristics of a and s used in the utility model, and $\xi_{as} \sim_{iid} EV(I)$ (logit) with usual variance normalization to $\pi^2/6$. (Results from a probit version are similar.) $\boldsymbol{W}_{as}$ contains additional factors that may affect reapplication acceptance decisions by NYC: a measure of how oversubscribed or popular s is (the number of applicants rejected by s in the initial application process), an indicator that a ranks s in the initial application, and another indicator that a is rejected by s in the initial application process. Let p_{as_i} be the estimate of $\Pr(a \text{ is re-assigned to } s_i)$ I obtain by applying the above model to the reapplication acceptance data.

The second step consists of applicants' reapplication decisions given acceptance probabilities in the first step. I assume that applicant a reapplies if there is a combination of schools (s_1, s_2, s_3) such that (a) for all $i = 1, 2, 3$, s_i is a school other than s_a^0 or empty, (b) $s_i \neq s_j$ or $s_i = s_j = \emptyset$ for all $i \neq j$, and (c)

$$\begin{aligned}
&\underbrace{\Sigma_{i=1}^3 p_{as_i} \bar{U}_{as_i}^1 + (1 - \Sigma_{i=1}^3 p_{as_i}) \bar{U}_{as_a^0}^1 - \bar{U}_{as_a^0}^1}_{\text{expected benefit from reapplying}} > \underbrace{c_a}_{\text{cost}} \\
&\Leftrightarrow \Sigma_{i=1}^3 p_{as_i} U_{as_i}^1 - \Sigma_{i=1}^3 p_{as_i} U_{as_a^0}^1 > c_a + \gamma_a \Sigma_{i=1}^3 p_{as_i} \equiv \bar{c}_a \\
&\Leftrightarrow \Sigma_{i=1}^3 p_{as_i} U_{as_i}^1 > \Sigma_{i=1}^3 p_{as_i} U_{as_a^0}^1 + \bar{c}_a, \qquad (3)
\end{aligned}$$

where c_a is the *reapplication cost*.[17] This model imposes the rational or sophisticated expectation assumption that each applicant believes that she is accepted by schools s_1, s_2, and s_3 with mutually exclusive probabilities p_{as_1}, p_{as_2}, and p_{as_3}, respectively. This assumption is unavoidable since there seems to be no way to identify subjective p_{as_i} for each (a, s_i) pair. The condition for reapplying can be written as in the last line of (3), a discrete choice with *switching costs* $\bar{c}_a$. I use "switching costs" to mean such combinations of reapplication costs and initial assignment effects, which I will separately identify.

If reapplying, applicant a ranks schools (s_1, s_2, s_3) to maximize the expected benefit in the left hand side of (3), i.e., schools with largest $p_{as} U_{as}^1$. I also assume that if applicant a

[17] A small fraction of applicants are not assigned to any school in the initial match. The above model is not well-defined for these unassigned applicants since $s_a^0 = \emptyset$ for them. For them, I assume the following model for the utility of the outside option $\emptyset$: $U_{a\emptyset}^1 = U_\emptyset^1 + \epsilon_{a\emptyset}$ where $U_\emptyset^1$ is the outside-option-specific constant and $\epsilon_{a\emptyset}$ is an unobserved utility shock. I make the same assumption for the second model below.

reapplies but does not exhaust her new preference list, i.e., a ranks less than three schools in $\succ_a^1$, any unranked school is less preferred to the guaranteed initial assignment s_a^0 in $\bar{U}_{as}^1$.

The above reapplication acceptance model has limitations. For example, ideally, I would let b_i be heterogeneous across schools or applicants. However, this is infeasible because there are more than 700 schools, while only about 6000 applicants reapply, and each reapplicant ranks at most three schools in the reapplication. Instead, I include rich attributes of applicants and schools in $\boldsymbol{X}_{as}$ and $\boldsymbol{W}_{as}$. I also have to exclude the effects of $s' \neq s$ on the probability that applicant a's reapplication for school s is accepted. If acceptance probability p_{as_i} depends not only on school s_i but also on s_j, then the maximizer of the expected benefit of reapplying is not necessarily schools with largest $p_{as}U_{as}^1$. To find the maximizer, I need to search over all possible combinations of up to 3 schools. Since there are more than 700 schools, the number of such combinations is prohibitively large, making estimation intractable. Due to these difficulties, I resort to the above simplified model.[18]

Model of Reapplications 2: Naive Free Expectation. The main concern with the rational expectation model is that it imposes rational expectations. To deal with this issue, I also consider an alternative model that does not assume rational expectations, but instead assumes naive beliefs about how the reapplication process works. In the alternative model, applicant a does not reapply if

$$\begin{aligned}
&\underbrace{p_a(\max_{s\neq s_a^0}\bar{U}_{as}^1 - \bar{U}_{as_a^0}^1)}_{\text{expected benefit from reapplying}} < \underbrace{c_a}_{\text{cost}}(>0) \\
&\Leftrightarrow \max_{s\neq s_a^0} U_{as}^1 - U_{as_a^0}^1 < c_a/p_a + \gamma_a \equiv \tilde{c}_a \\
&\Leftrightarrow U_{as_a^0}^1 + \tilde{c}_a > U_{as}^1 \text{ for any } s \neq s_a^0, \qquad (4)
\end{aligned}$$

where c_a is the reapplication cost and p_a is a's subjective probability that a's reapplication is accepted. I do not assume p_a to be the same as the real reapplication acceptance probability.

This model imposes a simplifying assumption that the expected benefit from reapplying is expressed as p_a times the utility difference between the initial assignment and the new most preferred school. If s_a^0 is a's most preferred school in $\bar{U}_{as}^1$, i.e., $\max_s \bar{U}_{as}^1 = \bar{U}_{as_a^0}^1$, then the left hand side of the second line of (4) is negative and so a never reapplies. Otherwise, a reapplies when the expected benefit from doing so exceeds the reapplication cost c_a. Again, the condition can be written as in the last line of (4), a discrete choice with switching costs $\tilde{c}_a$.

[18] In general, the above model can cause internal inconsistencies $\Sigma_{i=1}^3 p_{as_i} > 1$. In my data, however, reapplication acceptances are rare, and estimated p_{as_i} is almost always less than 0.2. As a result, $\Sigma_{i=1}^3 p_{as_i} > 1$ never happens.

For those who reapply, I observe new rank-ordered preference $\succ_a^1$ and assume that each reapplicant submits $\succ_a^1$ based on $\bar{U}_{as}^1$'s, i.e., $s \succ_a^1 s'$ only if $\bar{U}_{as}^1 > \bar{U}_{as'}^1$. As in the rational expectation model, I also assume that if a reapplies but does not exhaust her new preference list, i.e., applicant a ranks less than three schools in $\succ_a^1$, any unranked school is less preferred to the guaranteed initial assignment s_a^0 in $\bar{U}_{as}^1$.

Comparing the two models, the rational expectation model allows for school-specific reapplication acceptance probabilities, while assuming rational expectation. The naive free expectation model does not need the rational expectation assumption, but assumes that reapplicants form simplistic beliefs about how the reapplication process works. These two models are thus expected to be complementary and serve as robustness checks for each other. I estimate both models and show that the key results hold under both models. Before moving on to identification and estimation, however, I need to discuss other important modeling decisions.

Discussions of Modeling Assumptions

Choice Frictions and Learning. The key modeling decision about the evolving utility model is how to model frictions in initial choices. My model specifies them as $\beta_{ak}(1+f_{ak})X_{ask}$. Ideally, I would like to make frictions f_{ak} more flexible, for example, f_{ask} that is heterogeneous not only across applicants a and characteristics k, but also across schools s. Alternatively, an additive specification $\beta_{ak}(X_{ask} + f_{ask})$ may be another more flexible way to model frictions. In such more flexible models, $\beta_{ak}(1 + f_{ask})$ or $\beta_{ak}(1 + \frac{f_{ask}}{X_{ask}})$ (note that $\beta_{ak}(X_{ask} + f_{ask}) = \beta_{ak}(1 + \frac{f_{ask}}{X_{ask}})X_{ask}$) performs the same role as the taste coefficient on X_{ask} in usual discrete choice models with no frictions. However, it is unclear how I can identify the distribution of such coefficients that depend not only on a but also on s. For example, consider a static discrete choice model with no frictions $U_{as} = \beta_{as}d_{as} + \epsilon_{as}$ where d_{as} is the distance, which is always positive. For each a, consider any preference coefficients $(\beta_{as})_s$ such that (1) each β_{as} is so large that the effect of ϵ_{as} on choice probabilities is negligible and (2) $\beta_{as_1} > \beta_{as_2} > ...$, where s_i is a's observed i-th choice. Any such $(\beta_{as})_s$ can rationalize a's observed choices, making identification impossible without a particular parametric assumption. To avoid the potential lack of identification, I make f_{ak} independent from s.[19] Even under this restriction, f_{ak} allows for rich heterogeneity across a and k.

[19]This discussion relates to recent attempts to identify and estimate discrete choice models with measurement errors, since frictions about X_{ask} or β_{ak} can be reinterpreted as measurement errors in X_{ask}. See Hu (2015) for a review.

Yet another potential way to model frictions and learning is to introduce "consideration sets", i.e., subsets of schools applicants consider when they make initial choices. See Goeree (2008) for an existing empirical model of consideration sets. There is not enough variation in my data to allow for both frictions f_{ak} and consideration sets. Given a choice between frictions and consideration sets, I prefer frictions for several reasons. First, consistent with the friction specification, self-reported reasons for reapplications mention the initial lack of knowledge about school characteristics or their preferences about school characteristics more often than the initial lack of knowledge about the presence of particular schools. Second, for inferring a consideration-set-formation process, I need some variation that makes different schools more or less likely to enter consideration sets. However, it is unclear if time-series variation in my data is enough, since the contrast between initial applications and reapplications contains no variation across schools. For these reasons, this paper focuses on frictions f_{ak}, and I leave a consideration-set approach for future research.

Deliberate Reapplication Decisions. My reapplication model assumes that families make reapplication decisions by deliberately comparing their initial assignments with other schools according to their new demand. A potential concern is that reapplication decisions may be primarily driven by less systematic factors (e.g., inattention unrelated to initial assignments). However, there are several descriptive facts showing that, consistent with my model, families' reapplication behavior responds to the desirability of their initial assignments. For example, the more preferred school an applicant is initially assigned, the less likely she is to reapply (Figure 4); this correlation is always present across many subgroups defined by demographic characteristics and first-round application behavior (Appendix Figure A.5). The next identification section will detail these facts. Appendix A.3.1 further shows that this correlation is causal and structural. This suggests that many applicants, including those who do *not* reapply, compare their initial assignments and other schools.

In addition, the amount of choice reversals reapplicants exhibit is strongly correlated with the preference rank of their initially assigned school (Figure 3). These correlations are also consistent with my model. In my framework, the more preferred school an applicant is initially assigned, the smaller the expected benefit of reapplying is for her. Thus, those assigned to more preferred initial assignments need to experience larger demand changes to find it worth reapplying, compared with those assigned to less preferred initial assignments. As a result, conditional on reapplying, the amount of observed choice reversals, which reflect demand changes, should be decreasing in the preference rank of the initially assigned school, implying a pattern as in Figure 3.

These facts suggest that many families behave in ways consistent with deliberate reap-

plication decisions. If some people behave according to inattention unrelated to initial assignments, their behavior is likely to be absorbed by reapplication costs in my model. This misspecification concern is common in empirical studies on switching costs.

Truthful Behavior in Initial Applications. The model assumes that each applicant makes the initial preference $\succ_a^0$ as a non-strategic rank-ordered discrete choice based on old utilities U_{as}^0's. This should be a reasonable assumption since (1) a majority of applicants (more than 70%) do not exhaust preference lists and rank 11 or fewer schools, and (2) the deferred acceptance algorithm used in the initial application process is strategy-proof for applicants and guarantees that the above truthful behavior is always optimal for any applicant who does not exhaust her preference list. Even for those who exhaust their preference lists, the deferred acceptance algorithm makes it always optimal for any applicant to truthfully report her relative preference order over ranked schools.

The above discussion ignores the presence of the reapplication process. In principle, applicants may strategize in initial applications for switching to a more preferred school in the reapplication process. However, such strategic behavior is unlikely to benefit reapplicants: The reapplication acceptance rate is low (21%), and it is rare that reapplicants can switch to more preferred schools. Also, the reapplication process is a discretionary process with no algorithmic rule. It is thus unclear how to strategize in initial applications to benefit in the reapplication process. For example, one may suspect that in the reapplication process, it may be easier to be transferred to a school that is not ranked in initial applications, making it profitable to strategically drop some schools from initial applications. However, in the data, reapplication acceptances are more likely to be given to schools ranked in the first round.

Finally, there is an additional tractability consideration that forces me to ignore potential strategic behavior. With strategic behavior, I need to consider applicants' choices over *combinations (lists) of* schools, but the number of such combinations is prohibitively large in my setting with hundreds of schools. In light of these computational and conceptual reasons, I assume away strategic behavior in initial applications. Existing studies also use similar truth-telling assumptions (Hastings et al., 2008; Ajayi, 2013; Abdulkadiroğlu et al., 2015).

Outside Option. I do not explicitly model the outside option because it is unclear whether the data is informative about the outside option. Many applicants do not exhaust their initial preference lists, but it does not necessarily mean that all unranked schools, which are almost all of NYC schools, are less preferred over the outside option for them. A more reasonable interpretation seems to be that they are optimistic and expect that they will be assigned to

one of the ranked schools for sure; see Robbins (2011) for an article that reports about such optimistic families. In this scenario, the data does not provide any information about the comparison between the outside option and NYC schools. I thus refrain from modeling the outside option.

3.2 Identification

My model allows for preferences ($\boldsymbol{\beta}_a$), frictions in initial choices ($\mathbf{f}_a$), reapplication costs (c_a), and initial assignment effects (γ_a). This flexibility may create an identification concern, because it is often difficult to separately identify heterogenous preferences and switching costs (Chamberlain, 1983; Heckman, 1991). In fact, no other model reviewed in the related literature seems to allow for all of the above model components simultaneously. This section explains which aspects of my data allow me to distinguish the model components. For brevity, I focus on the above key parameters and ignore school-specific effects U_s^t and unobserved utility shocks ϵ_{as}^t.

By identification results for standard discrete choice models with no frictions (Matzkin, 2007; Manski, 2009; Berry and Haile, 2010; Fox et al., 2012), the data on initial preferences $\succ_a^0$ identifies the distribution of $\beta_{ak}(1+f_{ak})$: It is because in initial utilities U_{as}^0, the composite term $\beta_{ak}(1+f_{ak})$ has the same role as the preference coefficient in standard models. [20] However, it is not possible to separate out preferences $\boldsymbol{\beta}_a$ and frictions $\mathbf{f}_a$ by the initial application data alone: insensitivity of choices to a certain characteristic (e.g., academic performance) may be because of weak preferences for it or frictions about it.

I need to use the data on reapplications to distinguish frictions $\mathbf{f}_a$ and preferences $\boldsymbol{\beta}_a$. Let me use the term *demand changes* to mean utility changes by frictions about characteristics $\mathbf{X}_{as}$, i.e., $f_{ak}\beta_{ak}X_{ask}$'s. In combination with already-identified $\beta_{ak}(1+f_{ak})$, identification of demand changes is enough to separate out frictions $\mathbf{f}_a$ and preferences $\boldsymbol{\beta}_a$. The identification of demand changes is aided by the increase in the correlation between characteristics $\mathbf{X}_{as}$ and choices induced by U_{as}^t from $t=0$ to $t=1$ (recall Tables 3 and 4). The difficulty is that reapplication behavior is subject to switching costs ($\bar{c}_a$ in the rational expectation model and $\tilde{c}_a$ in the simplistic free expectation model). It is usually hard to separate demand changes from switching costs, because the low rate of choice changes, which we observe in

[20] My empirical setting has an additional advantage in that the initial application process is the first time for applicants to choose among schools in the system. This means that the initial application process is the initial period of the dynamic school choice process I try to model. Initial applications $\succ_a^0$ are *not* subject to any switching costs or state dependence stemming from prior choices. As a result, I do not need to be concerned about the econometric initial condition problem, which plagues usual dynamic choice models where the econometrician does not observe initial choices (Chamberlain, 1983). This enables the identification of $\beta_{ak}(1+f_{ak})$.

most settings including mine, is usually consistent both with small demand changes and large switching costs (Chamberlain, 1983; Heckman, 1991). The main identification challenge is thus how to separately identify demand changes and switching costs. Moreover, I need to resolve an additional difficulty of how to decompose switching costs into reapplication costs c_a and initial assignment effects γ_a.[21]

To overcome these challenges, I exploit the following three institutional features of centralized school choice systems, including NYC's.

(1) *Capacity constraints.* Due to capacity constraints, many applicants are initially assigned to a school other than the most preferred school.

(2) *Partially random initial assignments.* In the first-round assignment mechanism, to determine initially assigned schools s_a^0, the algorithm uses student and school preferences, but school preferences are coarse or weak in that a school's preference is indifferent among many students. NYC draws random lottery numbers to break ties or indifferences, and uses the resulting strict school preferences to compute initial assignments. This use of admissions lotteries makes initial assignments partially random.

(3) *Rank-ordered reapplication preferences.* Reapplicants make new, rank-ordered school choices in their reapplications, as explained in section 2.1.

Demand Changes and Switching Costs. Intuitively, the identification logic is as follows. For simplicity, let me ignore demand responses to initial assignments and focus on demand changes and reapplication costs. Assume that there are only two preference ranks, the first choice and the lower choice. Many applicants are "lower-choice non-reapplicants," who are initially assigned to their old lower choice but do not reapply, due either to reapplication costs or demand changes. This fact allows me to measure the total effects of demand changes and reapplication costs. There are also "first-choice reapplicants," who are initially assigned to their old first choice but reapply, which must be due to demand changes. Other applicants assigned to their first choice may also experience demand changes but are locked in by reapplication costs. The fraction of first-choice reapplicants in all applicants assigned to the first choice thus tells me the difference between the effects of demand changes and switching costs.

Now suppose that admissions lotteries in the first-round mechanism guarantee that initial assignments are randomly assigned and that applicants assigned to the first-choice and

[21] The rational expectation model of reapplications involves an additional complication that reapplication decisions and preferences are based on utilities $\bar{U}_{as}^1$ weighted by reapplication acceptance probabilities p_{as}. Identification results for discrete choice models with similar complications appear in, e.g., Agarwal and Somaini (2015).

lower-choice are comparable people with similar demand changes and reapplication costs. I can compare the fractions of lower-choice non-reapplicants and first-choice reapplicants to measure the amount of reapplication costs. Heuristically,

(fraction of non-reapplicants among applicants assigned to the initial lower choice)
-(fraction of reapplicants among applicants assigned to the initial first choice)
=(demand changes+reapplication costs)-(demand changes-reapplication costs)
=2×reapplication costs,

which separates reapplication costs from demand changes.

More precisely, thanks to institutional feature 1 (capacity constraints), I can compute moments like Figure 4a, where the solid black line relates the preference rank (with respect to initial $\succ_a^0$) of the initially assigned school s_a^0 to the conditional probability of reapplying. The line is upward-sloping, i.e., the more preferred school an applicant is assigned, the less likely she is to reapply.

This moment turns out to contain information to separate demand changes and switching costs. To illustrate this, let me start with the question of what the line should look like if there were no demand changes ($\Sigma_{k=1}^{K}\beta_{ak}f_{ak}X_{ask} = 0$) and no switching costs ($\bar{c}_a = 0$ or $\tilde{c}_a = 0$). The answer is the dotted red line in Figure 4a: If an applicant is assigned the old first choice, it remains to be the new first choice by the no-demand-change assumption, and there is no reason for her to reapply. If an applicant is assigned an old lower choice, by the no-demand-change assumption, the initially assigned school remains to be different from the new first choice. Thus, she should reapply as long as there is no switching cost, since reapplying gives her a positive probability of being re-assigned to a more preferred school without risking the initial assignment. (The initial assignment is guaranteed even if an applicant reapplies, as explained in Section 2.1.)[22] This hypothetical dotted line under no demand changes and no switching costs is different from the real solid line. The discrepancy between the two has to be caused by demand changes *or* switching costs.

To distinguish demand changes and switching costs, I use institutional feature 2 (partially random initial assignments). For simplicity, start by assuming that initial assignments are purely random. After illustrating the identification logic in this simplified case, I explain how to extend the identification logic to the case with real partially random initial assignments. Let me consider the question of what the black solid line should look like under no switching

[22]This reasoning assumes that the probability of a reapplication acceptance is positive for all possible preference ranks of the initial assignment. Appendix Figure A.4 confirms that this is the case in the data. Estimated p_{as_i} for the rational expectation model is also always positive.

costs but with potential demand changes. To answer this, in Figure 4b, consider the solid short red arrow above "1" on the x axis. Let R be the length of the solid red arrow.

$$
\begin{aligned}
& R \\
& \equiv \Pr(a \text{ reapplies} | s_a^0 \text{ is } a\text{'s old 1st choice}) \\
& = \Sigma_{K \geq 2} \Pr(a\text{'s old } K\text{-th choice}{=}a\text{'s new 1st choice} |\ s_a^0 \text{ is } a\text{'s old 1st choice}) \\
& \quad \text{(by the assumption that there are no switching costs)} \\
& = \Sigma_{K \geq 2} \Pr(a\text{'s old } K\text{-th choice}{=}a\text{'s new 1st choice} |\ s_a^0 \text{ is } a\text{'s old } K\text{-th choice}) \\
& \quad \text{(by random assignment of } s_a^0)^{23} \\
& = \Sigma_{K \geq 2} \Pr(a \text{ does not reapply} |\ s_a^0 \text{ is } a\text{'s old } K\text{-th choice}) \\
& \quad \text{(by the assumption that there are no switching costs)} \\
& \equiv B,
\end{aligned}
$$

where B is defined as the sum of the lengths of the dotted blue arrows in the figure. However, Figure 4b shows that the sum of the lengths of the dotted blue arrows B (9.510) is larger than the length of the solid red arrow R (0.025). This means that, regardless of the amount of demand changes, the no-switching-cost assumption leads to a contradiction with the data. In contrast, if there are switching costs ($\tilde{c}_a > 0$), the model's requirement changes to

$$
\begin{aligned}
& R \\
& < \Sigma_{K \geq 2} \Pr(a\text{'s old } K\text{-th choice}{=}a\text{'s new 1st choice} |\ s_a^0 \text{ is } a\text{'s old 1st choice}) \\
& < \Sigma_{K \geq 2} \Pr(a \text{ does not reapply} |\ s_a^0 \text{ is } a\text{'s old } K\text{-th choice}) \\
& \equiv B,
\end{aligned}
$$

where previous equalities change to inequalities because additional people stop reapplying because of positive switching costs. The inequalities are consistent with Figure 4b. In this way, the discrepancy between the solid red arrow R (the probability of reapplying conditional on being initially assigned the old first choice) and the dotted blue arrows B (the sum of the

[23] This step assumes that the preference rank of the randomly assigned initial assignment s_a^0 does not have direct effects on demand changes $\beta_{ak} f_{ak}$. Analogous assumptions are implicitly made in existing studies of switching costs reviewed in the related literature section in the following sense. In general, the separation of switching costs from heterogenous preferences needs two assumptions: (i) The status quo alternative as the source of switching costs is (at least partly or conditionally) randomly assigned. (ii) The status quo does not have direct effects on any demand or preference changes. Existing studies use the contrast between active vs passive choice periods or variation in supply-side advertisements to defend (i). In addition, they need (ii) and usually assume that all demand changes are exogenous. By contrast, my approach as outlined above uses a really random assignment to guarantee (i). In addition, while I need to assume that the preference rank of the randomly assigned initial assignment s_a^0 does not have direct effects on $\beta_{ak} f_{ak}$, I allow for a certain direct effect of s_a^0 on new demand through initial assignment effects $\gamma_a 1\{s = s_a^0\}$.

probabilities of *not* reapplying conditional on being assigned to old non-first choices) has to be driven by switching costs and not by demand changes, telling us the amount of switching costs.

Figures 4c and 4d summarize the separate identification of switching costs and demand changes. Starting from the solid black line, move it up until the point where the solid red arrow and the dotted blue arrows are balanced. The new dotted blue line with triangle markers describes such a point. The difference between the solid black line and the dotted blue line with triangle markers has to be due to switching costs, while the remaining difference between the two dotted lines is due to demand changes. Figure 4d thus suggests both significant demand changes and significant switching costs.

The above discussion shows that the probability of reapplying conditional on the preference rank of the initial assignment reveals the amount of demand changes and switching costs. Appendix Figures A.5a and A.5b use this logic to suggest that different demographic groups face different amounts of demand changes and switching costs. Racial minorities and academically struggling families exhibit flatter gradients, a sign of larger demand changes by learning.[24] Given this, I let the distribution of demand changes and switching costs as well as the other parameters heterogeneous across demographic groups in the estimation. In contrast, I find little heterogeneity among students with different initial assignments or different types of initial application behavior (Appendix Figures A.5c and A.5d). This may suggest that conditional on the preference rank of the initially assigned school, its identity may not matter for the amount of switching costs and demand changes.

Handling Partial Randomization. The above identification logic is under the simplifying assumption that initial assignments are completely randomly assigned. In the NYC school choice system, however, initial assignments are not purely random since they are confounded by non-random preferences of applicants and their priorities at schools. Nevertheless, the above identification logic extends to the more complicated real case with partially random initial assignments.

NYC creates initial assignments via the deferred acceptance algorithm that uses many inputs such as preferences, priorities, lottery numbers, and capacities (see Section 2 and Appendix A.4). Depending on these factors, different applicants have different assignment probabilities at schools. Yet, there is a way to find a set of applicants who share the same assignment probability at any school. Let me refer to a student's entire preference list and priorities at all schools as her *type*. The deferred acceptance algorithm treats students of

[24] This sheds additional light on heterogeneity of school choice behavior across demographic groups. See Hastings et al. (2008) and Nathanson et al. (2013) among others for related findings.

the same type symmetrically in that everyone of a given type faces the same probability of assignment to any school. This is because the only information about a student the algorithm uses is her preference, priorities, and lottery number; conditional on type, therefore, all that remains to determine her assignment is her lottery number, which is independently and identically distributed across students.

This gives me a way to extend the identification analysis in Figure 4 to the case with partially random initial assignments: I can simply repeat the same analysis *conditional on each type* to separately identify switching costs and demand changes for that type. Switching costs and demand changes are thus allowed to be heterogenous across different types without sacrificing identification. As far as identification is concerned, partial randomization does not cause any serious problem. The empirical implementation of the identification argument in Figure 4d is also robust to the explicit consideration of imperfectly random initial assignments. Appendix Figure A.6 provides a structural/causal version of Figure 4d that incorporates partially random assignment.

Initial Assignment Effects and Reapplication Costs. Now that I have explained how to identify $\boldsymbol{\beta}_a$ (preferences), $\mathbf{f}_a$ (frictions in initial choices), and switching costs, the final step is to decompose switching costs into reapplication costs c_a and initial assignment effects γ_a. Conditional on reapplying, reapplication costs c_a are sunk and do not affect new preferences reported in reapplications. On the other hand, initial assignment effects γ_a remain to affect new preferences since they directly enter the utility of the initial assignment s_a^0. I use this fact and institutional feature 3 (rank-ordered reapplication preferences) to derive the following two restrictions on γ_a: (1) If applicant a reapplies to switch from initial assignment s_a^0 to s, then $U_s^1 > U_{s_a^0}^1 + \gamma_a$. (2) If applicant a reapplies but does not exhaust her new preference, i.e., rank only one or two schools, then $U_{s_a^0}^1 + \gamma_a > U_s^1$ for every unranked school s. These restrictions involve initial assignment effects γ_a but *not* reapplication costs c_a, allowing me to separate out initial assignment effects γ_a. New preferences contain a lot of information about initial assignment effects γ_a since every reapplicant prefers some schools over the initial assignment s_a^0, while about 40% of reapplicants do not exhaust new preferences, implying that they prefer the initial assignment s_a^0 over unranked schools.

3.3 Estimation

In the estimation, characteristics $\mathbf{X}_{as}$ include those frequently mentioned in reapplication reasons: the road distance between applicant a's and school s's locations, school s's academic performance, type, size, and age. See Appendix A.1 for the construction of these variables.

For computational tractability and finite sample statistical precision, I need to impose distributional assumptions common to empirical discrete choice models. Let g_a be applicant a's demographic group defined by whether a is white/asian or black/hispanic and whether a's grade 7 reading grade category is high/middle or low (four groups in total). I assume that

- the period 0 coefficient $\beta_{ak}(1+f_{ak})$ on X_{ask} is iid according to
$$\begin{cases} \log N(\mu_{0k}^{g_a}, \sigma_{0k}^{g_a}) \text{ for negative distance or high academic performance} \\ N(\mu_{0k}^{g_a}, \sigma_{0k}^{g_a}) \text{ for any other characteristic,} \end{cases}$$
- the change in the coefficient due to learning is $\beta_{ak} f_{ak} \sim_{iid} N(\mu_{1k}^{g_a}, \sigma_{1k}^{g_a})$,
- $U_s^0 \sim_{iid} N(0, \sigma_0^{g_a})$,
- $U_s \sim_{iid} N(0, \sigma_1^{g_a})$[25],
- $\gamma_a \sim_{iid} N(\mu_\gamma^{g_a}, \sigma_\gamma^{g_a})$,
- $c_a \sim_{iid}$truncated $N(\mu_c^{g_a}, \sigma_c^{g_a})$ in the rational expectation model, and
- $c_a/p_a \sim_{iid}$truncated $N(\mu_c^{g_a}, \sigma_c^{g_a})$ in the naive free expectation model.

Note that motivated by heterogeneity across demographic groups in key descriptive moments (Appendix Figure A.5 and Appendix Table A.1), the whole parameter vector is allowed to be heterogeneous across demographic groups. I assume $\epsilon_{as}^0, \epsilon_{as} \sim_{iid} EV(I)$ (logit) with usual variance normalization to $\pi^2/6$.[26]

These assumptions make the model within each period a random coefficient logit model, which allows for flexible substitution patterns among schools (Train 2009 chapter 6). Appendix A.3.2 derives a partly analytical joint likelihood function for a sequence of initial application preferences in $t=0$ to reapplication decisions and preferences in $t=1$. This likelihood function is parametrized by $\theta \equiv ((\mu_{0k}^g, \sigma_{0k}^g, \mu_{1k}^g, \sigma_{1k}^g)_{k=1,\ldots,K}, \sigma_0^g, \sigma_1^g, \mu_\gamma^g, \sigma_\gamma^g, \mu_c^g, \sigma_c^g)_{g \in G}$ where G is the set of the four demographic groups defined above.

I estimate parameter θ using maximum simulated likelihood with 400 simulations using scrambled randomized Halton draws (Train 2009 chapter 9). The number of simulations appears to be enough for convergence because the estimates change little from 200 simulations

[25] It is computationally prohibitive to separately estimate U_s^t for each school s since the sample contains a large number of schools. Instead, I adopt this random effects specification in the spirit of Rossi et al. (1996) and Abdulkadiroğlu et al. (2015).

[26] It is possible and more desirable to estimate the variance of ϵ_{as} rather than assuming it. Computational difficulties prevent me from reporting results from such an extension in this draft, however. I verified that the results reported below are robust (i.e., changes are less than 10%) to ignoring ϵ_{as} and using only the other parts of demand changes.

to 400 simulations. I compute standard errors using the information identity with the Hessian being estimated by the outer product of the gradient of the simulated likelihood at the estimated parameter $\hat{\theta}$ (Train 2009 chapter 9). I assume the utility function U^1_{as} to be quasi-linear in distance and will often measure results by distance-equivalent utilities in $t = 1$.

This estimation procedure uses admissions lotteries as follows. The likelihood involves initial assignments s^0_a's as arguments, and I substitute realized initial assignments s^0_a's in the data into the likelihood.[27] As explained in the last identification section, these initial assignments s^0_a's are conditionally randomly assigned by the first-round deferred acceptance algorithm with admissions lotteries; I provide empirical support for conditionally random assignment in Appendix A.3.1. As a result, consistent with the identification argument using randomness in initial assignments, the likelihood-based estimation procedure uses the fact that initial assignments are conditionally random. Additional details of the estimation procedure and the construction of the estimation sample are in Appendix A.3.2.

3.4 Estimates and Fit

The parameter vector θ are high dimensional and not easy to directly interpret. I summarize key features of the estimated $\hat{\theta}$ here. (Appendix Tables A.6-A.9 show that many dimensions of $\hat{\theta}$ are significant and exhibit sizable heterogeneity.) First, demand changes by learning are an important feature of my empirical model. Figure 5a plots the distribution of estimated demand changes due to learning about observable school characteristics, i.e., $\Sigma^K_{k=1}\beta_{ak}f_{ak}X_{ask}$, against estimated overall new utilities ($\hat{U}^1_{as}$; the black distribution).[28] This figure is based on the estimated rational expectation model. As can be expected from the largeness of NYC, overall utilities from schools are highly dispersed, and most and least preferred schools are different by more than 50 mile-equivalent units. If there were no frictions and $\hat{\mathbf{f}}_a = 0$, the distribution of demand changes would be degenerate at the origin. This is far from the case, and there are large demand changes, as shown in Figure 5a. Their magnitude is often equivalent to multiple miles, though as expected it is smaller compared with the dispersion of overall utilities.

Despite these demand changes, and despite the fact that a majority of applicants are initially assigned to a non-first-choice school, only a small fraction of applicants reapply. To explain this fact, Figure 5b plots the distribution of estimated initial assignment effects ($\hat{\gamma}_a$; the left blue distribution) and reapplication costs ($\hat{c}_a$; the left red distribution) by simulating

[27] I do not simulate initial assignments since applicants make reapplication decisions and preferences conditional on realized initial assignments s^0_a's in the data.

[28] Recall that $\text{Var}(\epsilon^1_{as})=\text{Var}(\epsilon^0_{as} + \epsilon_{as}) = 2\text{Var}(\epsilon^0_{as})$. When comparing period 0 and 1 utilities, I divide period 1 utilities by 2 so that the variance of the unobserved utility shock stays the same between $t = 0$ and $t = 1$.

the estimated rational expectation model. The figure compares them against estimated overall new utilities ($\hat{U}^1_{as}$; the right black distribution). Estimated reapplication costs $\hat{c}_a$ and initial assignment effects $\hat{\gamma}_a$ are often significant and positive: Applicants not only face reapplication costs, but they also come to prefer their initially assigned schools more, compared with other schools. Consequently, unless the initially assigned school turns out to be much worse than a more preferred school, applicants do not find reapplying worthwhile, which explains the low observed reapplication rate. Similar patterns for the alternative naive free expectation model are reported in Appendix Figures A.7a and A.7b.

These estimation results confirm the suggestive descriptive evidence (in Sections 2 and 3.2) that there are demand changes by learning, but switching costs (combinations of reapplication costs and initial assignment effects) prevent many demand changes from translating into observed choice changes. In addition, since the structural estimates show that both reapplication costs and initial assignment effects lower the reapplication rate, I can quantify the relative contributions of learning, reapplication costs, and initial assignment effects to reapplication behavior.

As an evaluation of how well the estimated model fits the data, I simulate the estimated model and key moments 50 times, and compare the average simulated moments against the real ones. Figure 6 plots the real first choice market shares of schools in the initial application process against the simulated ones, where the *first choice market share* of a school is defined as the fraction of applicants who rank it first among all applicants who make a first choice. The real and simulated shares are highly correlated. On the reapplication behavior, Table 5 compares the most important moments — the fractions of reapplicants and those who exhibit choice reversals — between the data and the estimated model. Both models mimic the behavior of the data well in terms of these moments.

Finally, Table 6 compares the moment in Table 4 — changes in R^2s from school-level regressions of schools' first choice market shares on observable school characteristics — between the data and the model. The model resembles the data in that reapplicants' choices become more correlated with and responsive to observable school characteristics from the first round to the reapplication process. These results suggest that the estimated models do decent jobs at matching key moments in the data.

4 Welfare Consequences of Evolving Demand

4.1 Costs of Ignoring Demand Changes

Estimates of the structural model reveal many more demand changes than are suggested by the low observed reapplication rates I showed in Section 2. To measure the amount of hidden demand changes, Table 5 shows the fractions of reapplicants as well as those with choice reversals in the counterfactual simulations of the estimated models without reapplication costs c_a (while keeping the estimated initial assignment effects $\hat{\gamma}_a$ as they are). With no reapplication costs, the reapplication rate increases from 7% in the data to 30-40%. More importantly, a majority of the counterfactual reapplicants exhibit choice reversals, implying that the fraction of applicants with underlying demand changes (more than 20%) is several times larger than the fraction of those with choice reversals in the data (5%). These demand changes are masked by reapplication costs.[29] This finding is consistent with the suggestive descriptive evidence of hidden demand changes in Figure 4.

As a result of these significant demand changes, the welfare cost of ignoring demand changes is large. To measure the welfare cost, I compare the real first-round assignment based on school choices induced by old demand U^0_{as} with the counterfactual "frictionless benchmark." The frictionless benchmark is defined as what would have been produced by the same first-round deferred acceptance algorithm, had families made choices based on their new demand U^1_{as}. Since the two differ only in whether families make choices based on old or new demand, the difference between the two captures the welfare costs of ignoring demand changes by learning.[30]

Table 7 summarizes welfare changes from the real first-round assignment to the frictionless benchmark.[31] The real first-round assignment and the frictionless benchmark are

[29] Initial assignment effects also lower the reapplication rate. As shown in Figure 5b, the estimates show that initial assignment effects are often positive and large, meaning that families tend to get to prefer initially assigned schools more, compared with other schools. This satisfaction with initial assignments lowers the reapplication rate by more than 20%, which is computed as the difference between the real reapplication rate and the counterfactual rate under no initial assignment effects $\gamma_a = 0$ (but with estimated reapplication costs $\hat{c}_a$).

[30] Except applicants' choices, every other input is unchanged between the real first-round assignment and the frictionless benchmark. For example, school capacities and their preferences or priorities over applicants are fixed.

[31] When simulating an assignment mechanism, I also simulate lottery numbers used by the mechanism to break ties in priorities. Recall the explanation in Section 3.2 about the use of lottery numbers in the NYC system, and see Appendix A.3.1 for further details. Furthermore, it is sometimes the case that an applicant's simulated preferences contain schools that the applicant does not rank in her real preference observed in the data. For such pairs of applicants and schools, the data does not provide priority information. In such cases, I simulate their priorities from the empirical distribution of priorities in the data. I do the same lottery number and priority simulation for the other counterfactual exercises. Finally, I assume that each assigned applicant experiences estimated initial assignment effect $\hat{\gamma}_a$ from her assignment. The results change little

significantly different. Specifically, the two assignments give different allocations (schools) to a majority of families. Also, the average welfare loss under the real first-round assignment compared with the frictionless benchmark is more than 1-mile-equivalent when I measure it by new demand U^1_{as} after learning, which is assumed to be quasi-linear in the distance between the family and the school locations. The 1-mile-equivalent utility unit can be interpreted as corresponding to traveling 1 mile every school day during the high school years. This magnitude corresponds to more than .15 standard deviation in the distribution of utilities from all schools for each applicant.[32] Significant learning and demand changes thus undermine the welfare performance of the initial match that ignores demand changes. This illustrates that demand-side choice frictions and learning significantly affect the welfare gains from a centralized market. This motivates me to investigate ways to accommodate demand changes, to which I turn next.

4.2 Evaluating Reapplication Processes

The large welfare costs of ignoring demand changes motivate me to investigate ways to alleviate the costs by accommodating demand changes. As already explained, NYC runs a discretionary, human-driven reapplication process, presumably to improve families' welfare by flexibly accommodating changing needs. It is also possible to run a centralized algorithm not only for the initial market but also for the reapplication process. This section evaluates how well the discretionary and counterfactual centralized reapplication processes alleviate the welfare cost of ignoring demand changes.

Theory

I start by introducing counterfactual centralized reapplication mechanisms. Appendix A.4 builds a dynamic version of the school-student assignment model, analyzes two centralized designs of the reapplication process, and shows that they are the "best possible" mechanisms to accommodate choice changes. The counterfactual mechanisms are what I call the *dynamic deferred acceptance mechanism* and the *deferred deferred acceptance mechanism*. They are easily implemented by applying the deferred acceptance algorithm to observed choice data.

To define the mechanisms, take as given applicants' old preferences $\succ^0_A \equiv (\succ^0_a)_a$ in initial application, their new preferences $\succ^1_A \equiv (\succ^1_a)_a$ (which I construct below), and school prefer-

even if I exclude initial assignment effects. The same treatment of initial assignment effects applies to the other counterfactual exercises.

[32] I use new demand U^1_{as} as my welfare measure since it is demand after leaning at a point in time closer to enrollment periods; new demand is thus expected to be a better welfare measure than old demand U^0_{as}. In other words, in this and other counterfactual welfare analyses, I assume that frictions $\mathbf{f}_a$ are welfare-*ir*relevant.

ences/priorities $\succ_S$ over applicants observed in the data. School preferences/priorities $\succ_S$ are strict preferences/priorities after tie-breaking by lottery numbers. The dynamic deferred acceptance mechanism determines an assignment as follows.[33]

(1) Compute the initial match $DA(\succ_A^0, \succ_S)$ in the initial application process, where function $DA(\cdot,\cdot)$ maps each possible profile of applicant and school preferences into the assignment produced by the deferred acceptance algorithm under these input preferences.

(2) Give an initial match guarantee to each applicant by modifying each school s's preference so that s most prefers applicants matched with s in step 1.

(3) Use modified school preferences/priorities $\succ_S'$ and applicants' reapplication preferences $\succ_A^1$ to get $DA(\succ_A^1, \succ_S') \equiv \varphi_{dynamic}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ where $\varphi_{dynamic}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ denotes the assignment under the dynamic deferred acceptance mechanism under input preferences $(\succ_A^0, \succ_A^1, \succ_S)$.

On the other hand, the deferred deferred acceptance mechanism is the same as the dynamic deferred acceptance mechanism except that I make no modification to school preferences/priorities. That is, $\varphi_{deferred}^{DA}(\succ_A^0, \succ_A^1, \succ_S) \equiv DA(\succ_A^1, \succ_S)$ where $\varphi_{deferred}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ denotes the assignment under the deferred deferred acceptance mechanism under input preferences $(\succ_A^0, \succ_A^1, \succ_S)$.

Appendix A.4 provides a formal analysis of $\varphi_{dynamic}^{DA}$ and $\varphi_{deferred}^{DA}$, and shows that they are the "best possible" mechanisms to accommodate choice changes. Specifically, consider the following criteria of how well a mechanism accommodates choices changes and caters to new preferences (Appendix A.4 formally defines these properties):

(I) "Fairness (stability)" with respect to $(\succ_A^1, \succ_S)$, i.e., no applicant-school pair "blocks" the outcome under that mechanism and has an incentive to jointly deviate from it to be matched with each other outside the mechanism.

(II) "Being less unfair (unstable)" than the initial match with respect to $(\succ_A^1, \succ_S)$, i.e., it is always the case that any applicant-school pair blocking the outcome under that mechanism also blocks the initial match.

(III) "Weak Pareto efficiency" with respect to $\succ_A^1$, i.e., there is no other assignment that every applicant strictly prefers over the outcome under that mechanism.

[33] Similar dynamic mechanisms have been discussed in existing studies (Pereyra, 2013; Coles et al., 2014; Kadam and Kotowski, 2015). The empirical and theoretical analysis in this paper does not appear to have an analog in their analyses.

(IV) "Always Pareto dominating the initial match" with respect to $\succ_A^1$.

(V) "Dynamic strategy-proofness," i.e., any preference manipulation by any applicant in any period is never strictly profitable with respect to that applicant's preference in that period.

In terms of these properties, $\varphi^{DA}_{dynamic}$ and $\varphi^{DA}_{deferred}$ are the best possible mechanisms, as the the following result (shown in Appendix A.4) implies. This result motivates using $\varphi^{DA}_{dynamic}$ and $\varphi^{DA}_{deferred}$ as the canonical centralized reapplication mechanisms.

Proposition 1 *1.A) $\varphi^{DA}_{dynamic}$ satisfies (II) being less unfair than the initial match, (III) weak Pareto efficiency, and (IV) always Pareto dominating the initial match (call this set of desiderata A), but not others.*
1.B) $\varphi^{DA}_{deferred}$ satisfies (I) fairness, (II) being less unfair than the initial match, (III) weak Pareto efficiency, and (V) dynamic strategy-proofness (call this set of desiderata B), but not others.
2) Consider any possible dynamic mechanism φ. φ can satisfy only a subset of set A or B.

Discretionary vs Centralized Reapplication Processes

My counterfactual analysis studies how well the counterfactual centralized reapplication processes improve on the initial match and alleviate the welfare costs of ignoring demand changes. I also compare the centralized processes with the existing discretionary process. There are two ways to implement this particular evaluation. First, I can evaluate the reapplication processes in a descriptive way, without resorting to the model. An alternative, usual approach is to evaluate each reapplication process by simulating the estimated model. I start with the descriptive evaluation and then compare it with the model-based evaluation.

For empirically implementing and evaluating counterfactual centralized mechanisms $\varphi^{DA}_{deferred}$ and $\varphi^{DA}_{dynamic}$, I need to define $\succ_a^1$, applicant preferences in the reapplication stage. In the descriptive counterfactual evaluation, I construct them as follows. If applicant a does not reapply in the data, then assume that $\succ_a^1$ stays the same at $\succ_a^0$ (the preference that applicant a submits in the initial application period). If a reapplies and reports a new reapplication preference, then define $\succ_a^1$ as a's reapplication preference followed by $\succ_a^0$. For example, if $\succ_a^0$ is (s_1, s_2) and a reapplies and ranks only s_3 in her reapplication preference, then $\succ_a^1$ is (s_3, s_1, s_2). This construction of $\succ_a^1$ ignores all unobserved demand changes and uses only observed reapplications and choice changes. Hidden demand changes (in Table 5) suggest that $\succ_a^1$ may not be an appropriate welfare measure for applicants who do not reapply. However, this is probably not a big problem for this particular evaluation of reapplication

processes, because reapplication processes mainly affect reapplicants, for whom $\succ_a^1$ reflects demand changes and is expected to be a reasonable welfare measure.[34] I use $\succ_A^0$ and $\succ_A^1$ created above (as well as school capacities and priorities in the data) to simulate $\varphi^{DA}_{dynamic}$, $\varphi^{DA}_{deferred}$, and the discretionary reapplication process, and I compare them with the initial match with respect to $\succ_A^1$. Note that this procedure makes no use of the empirical model and is based only on objects directly observed in the data.

The comparison between the discretionary reapplication process and the deferred and dynamic deferred acceptance mechanisms is in Figure 7. Starting from the initial match as the common status quo, each line plots the distribution of preference rank improvements (with respect to $\succ_a^1$) of the finally assigned school under each of the three reapplication processes. Figure 7 shows that all reapplication processes produce welfare gains. More importantly, the centralized reapplication processes are more effective and produce gains more than twice as large as those from the discretionary reapplication process.

Table 8 summarizes this result and compares it with the results from an alternative evaluation based on the estimated structural models. In the structural counterfactual evaluation, I use the estimated models to simulate old and new utilities U^0_{as}, U^1_{as}, reapplication costs c_a, initial assignment effects γ_a, and associated initial application preferences $\succ_A^0$ (over up to 12 schools) induced by U^0_{as}. I then substitute the simulated initial application preferences (as well as school capacities and priorities in the data) into the deferred acceptance algorithm to obtain the initial match. I use this initial match and simulated U^1_{as}, c_a, and γ_a to simulate reapplication decisions. If applicant a reapplies, let $\succ_a^1$ be the reapplication preference (over up to 3 schools) induced by U^1_{as}, followed by $\succ_a^0$ (as in the descriptive evaluation). Otherwise, define $\succ_a^1 = \succ_a^0$. I then use $\succ_A^1$ and $\succ_A^0$ to compute the allocations under $\varphi^{DA}_{deferred}$ and $\varphi^{DA}_{dynamic}$. I compare the allocations produced by $\varphi^{DA}_{deferred}$ and $\varphi^{DA}_{dynamic}$ with the initial match with respect to simulated U^1_{as} under each simulation, and take the average over 50 simulations.

The descriptive and structural evaluations show similar gains from the centralized reapplication processes, as shown in Table 8. This provides additional support for the finding that the centralized reapplication processes produce larger gains than those from the discretionary reapplication process. This also provides some confidence to the use of the estimated structural model for evaluating other counterfactual policies, for which no model-free or descriptive evaluation is possible. An example of a counterfactual policy that innately needs the estimated model is the frictionless benchmark (in Table 7) that accommodates all demand

[34] The discretionary reapplication process affects only reapplicants. Centralized reapplication processes influence not only reapplicants but also others, but Appendix Figure A.8 shows that most of applicants affected by centralized reapplication processes are reapplicants.

changes, many of which are unobserved (as shown in Table 5).

Reapplication Processes and Demand-side Inertia

The centralized reapplication processes are shown to accommodate observed reapplications and choice changes better than the discretionary reapplication process does. This evaluation of the centralized reapplication processes takes estimated reapplication costs as given. On the other hand, there are technological changes and school districts' and social entrepreneurs' initiatives that may ease reapplication costs (e.g., online systems for more easily making and updating school choices).

To measure the potential effects of such demand-side interventions, I finally investigate the performance of the centralized reapplication processes relative to their "frictionless implementation." The frictionless implementation of the mechanisms is the hypothetical, possibly infeasible implementation that turns off reapplication costs c_a so that applicants express demand changes with no barrier. The frictionless implementation is similar to the frictionless benchmark in Table 7. In particular, the frictionless implementation of $\varphi^{DA}_{deferred}$ is the same as the frictionless benchmark. The frictionless implementation is not subject to demand-side inertia due to reapplication costs, while the feasible centralized reapplication processes are; the difference between the two thus captures the welfare effect of demand-side inertia (given the market design fixed). For simulating and evaluating the frictionless implementation, I need to use the estimated empirical model. The use of the estimated model for evaluating the frictionless implementation is the same as that for the feasible implementation in Table 8, except that I set reapplication costs to $c_a = 0$ when simulating the model to evaluate the frictionless implementation.

The welfare effect of demand-side inertia is large, and the gains from the mechanisms change by several times depending on the extent of demand-side inertia caused by reapplication costs: Table 9 shows that starting from the counterfactual scenario with no switching costs, estimated switching costs dilute the gains from the mechanisms by more than 50%.[35] Because of inertia and the resulting low participation, the reapplication processes reveal only a small fraction of the demand changes families experience as they learn. As a result, the centralized reapplication processes achieve no more than 30% of the welfare gains from their frictionless implementation, which responds to all the demand changes families would

[35] In Tables 8 and 9, welfare calculations do not include reapplication costs c_a (but include initial assignment effects γ_a). This choice is in order to make the comparison in Table 9 conservative and nontrivial: if welfare calculations include c_a, then turning off c_a would trivially result in welfare improvements. Incorporating c_a does not change the qualitative comparison among reapplication systems in Tables 8 and 9 since reapplicants experience reapplication costs under any reapplication system, and the welfare loss from reapplication costs is cancelled out when I compare different reapplication systems.

express if there were no demand-side inertia. This suggests that the gains from a dynamic centralized market depend to a large extent on demand-side inertia that prevents families from reapplying and expressing demand changes.

5 Conclusion and Future Directions

Centralized market-like institutions have become a widespread form of public policy. The success of these markets hinges on the assumption that participants make well-informed choices upfront. In this paper, I use data from NYC's school choice system to evaluate this assumption. I show that, contrary to the premise of well-informed upfront choices, families' choices change after the initial match as they learn about schools. To recover underlying evolving demand, I develop an empirical model of evolving demand for schools under learning, reapplication costs, and initial assignment effects. I exploit institutional features of centralized school choice systems, especially admissions lotteries, to separately identify these model components.

The estimates suggest that there are substantially more changes in underlying demand than in observed choices. These significant demand changes undermine the welfare performance of the initial match, and result in large welfare costs of ignoring demand changes. The large welfare costs of ignoring demand changes motivate me to investigate dynamic mechanisms, which I show can best accommodate choice changes in theory. I empirically find that these mechanisms significantly improve on the existing discretionary reapplication process and initial match. Also, the gains from the mechanisms change greatly depending on the extent of demand-side inertia caused by reapplication costs. Thus, demand-side frictions (such as learning, demand changes, and inertia) affect the gains from a centralized market as much as its design.

The suggested importance of frictions in participants' choices opens the door to many empirical and methodological questions. An implication of my results is the potential importance of demand-side technological changes or policy interventions that may alleviate demand-side frictions (e.g., online systems for more easily making and updating school choices, applications for more easily searching school characteristics). Little is known about the effects of such demand-side interventions, aside from what Hastings and Weinstein (2008) and Andrabi et al. (2015) report. Another related question is about the relationship between dynamic choice behavior and subsequent outcomes. For example, few papers study whether later, presumably better-informed school choices result in changes in academic achievement and other behavioral outcomes.

Methodologically, this paper contributes to understanding families' school choice dynam-

ics. A variety of extensions are both possible and desirable. For example, extending the data and model to more than two periods would make it possible to study the speed of learning and the dynamics of inertia. Another potentially fruitful direction is to incorporate richer aspects of learning, e.g., latent consideration sets families use when making choices, families' anticipation and sophistication about future learning. Finally, while this paper focuses on welfare analysis based on assignments or offers, it is probably more desirable to study welfare from enrollment (rather than assignments as intermediate steps toward enrollment and educational experience). I leave these challenging directions for future research.

References

Abaluck, Jason and Jonathan Gruber, "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *American Economic Review*, 2011, *101(4)*, 1180–1210.

_ and _, "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance," 2013. Working Paper.

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School Choice: A Mechanism Design Approach," *American Economic Review*, 2003, *93*, 729–747.

_, Nikhil Agarwal, and Parag Pathak, "The Welfare Effects of Coordinated School Assignment: Evidence from the NYC High School Match," 2015. Working Paper.

_, Parag A. Pathak, and Alvin E. Roth, "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 2009, *99* (5), 1954–1978.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," 2015. Working Paper.

Ajayi, Kehinde F., "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," 2013. Working Paper.

Akbarpour, Mohammad, Shengwu Li, and Shayan Oveis Gharan, "Dynamic Matching Market Design," 2015. Working Paper.

Anderson, Ross, Itai Ashlagi, David Gamarnik, and Yash Kanoria, "A Dynamic Model of Barter Exchange," in "Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms" SIAM 2015, pp. 1925–1933.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," 2015. Working Paper.

Autor, David, Nichole Maestes, Kathleen Mullen, and Alexander Strand, "Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants," 2015. Working Paper.

Baccara, Mariagiovanna, SangMok Lee, and Leeat Yariv, "Optimal Dynamic Matching," 2015. Working Paper.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 2007, *115(4)*, 588–638.

Berry, Steven and Philip Haile, "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," 2010. Working Paper.

Blundell, Richard, "Consumer Behaviour: Theory and Empirical Evidence — a Survey," *Economic Journal*, 1988, *98*, 16–65.

Bronnenberg, Bart, Jean-Pierre Dube, and Matthew Gentzkow, "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 2012, *102* (6), 2472–2508.

Card, David and Dean R Hyslop, "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers," *Econometrica*, 2005, *73* (6), 1723–1770.

Casalmiglia, Caterina, Chao Fu, and Maia Güell, "Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives," 2014. Working Paper.

Chamberlain, Gary, "Panel Data," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics, Volume 2*, Oxford: North Holland, 1983.

Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 2012, *80*, 969–1018.

Ching, Andrew T., Tulin Erdem, and Michael P. Keane, "Learning Models: An Assessment of Progress, Challenges, and New Developments," *Marketing Science*, 2013, *32(6)*, 913–938.

Choi, Syngjoo, Shachar Kariv, Wieland Muller, and Dan Silverman, "Who Is (More) Rational?," *American Economic Review*, 2014, *104* (6), 1518–50.

Coles, Peter A, Clayton R Featherstone, John William Hatfield, Fuhito Kojima, Scott Duke Kominers, Muriel Niederle, Parag A Pathak, and Alvin E Roth, "Comment on the NRMP's "Supplemental Offer and Acceptance Program" Proposed to Replace the Post-Match Scramble," 2014. Working Paper.

Dinerstein, Michael and Troy Smith, "Quantifying the Supply Response of Private Schools to Public Policies," 2015. Working Paper.

Dube, Jean-Pierre, Gunter J. Hitsch, and Peter E. Rossi, "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, 2009, *46(4)*, 435–445.

Dur, Umut and Onur Kesten, "Sequential versus Simultaneous Assignment Processes and Two Applications," 2014. Working Paper.

Echenique, Federico, Sangmok Lee, and Matthew Shum, "The Money Pump as a Measure of Revealed Preference Violation," *Journal of Political Economy*, 2011, *119*, 1201–1223.

Epple, Dennis, Richard Romano, and Holger Sieg, "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education," *Econometrica*, 2006, *74* (4), 885–928.

Ergin, Haluk, "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 2002, *88*, 485–494.

Fang, Hanming, Michael P. Keane, and Dan Silverman, "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, 2008, *116* (2), 303–350.

Farber, Henry S. and Robert Gibbons, "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 1996, *111* (4), 1007–1047.

Farrell, Joseph and Paul Klemperer, "Coordination and Lock-In: Competition with Switching Costs and Network Effects," in Mark Armstrong and Robert H. Porter, eds., *Handbook of Industrial Organization, Volume 3*, Oxford: North Holland, 2007.

Fox, Jeremy T, Kyoo il Kim, Stephen P Ryan, and Patrick Bajari, "The Random Coefficients Logit Model is Identified," *Journal of Econometrics*, 2012, *166* (2), 204–212.

Goeree, Michelle Sovinsky, "Limited Information and Advertising in the US Personal Computer Industry," *Econometrica*, 2008, *76* (5), 1017–1074.

Handel, Benjamin R., "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 2013, *103*, 2643–2682.

_ **and Jonathan T. Kolstad**, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015, *105(8)*, 2449–2500.

Hastings, Justine and J.M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

_ , Christopher A. Neilson, and Seth D. Zimmerman, "The Effects of Earnings Disclosure on College Enrollment Decisions," 2015. Working Paper.

_ , T. Kane, and D. Staiger, "Heterogeneous Preferences and the Efficacy of Public School Choice," 2008. Working Paper.

He, Yinghua, "Gaming the Boston School Choice Mechanism in Beijing," 2014. Working Paper.

Heckman, James J., "Identifying the Hand of Past: Distinguishing State Dependence from Heterogeneity," *American Economic Review Papers and Proceedings*, 1991, *81(2)*, 75–79.

Ho, Kate, Joseph Hogan, and Fiona Scott Morton, "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program," 2015. Working Paper.

Hortacsu, Ali, Seyed Ali Madanizadeh, and Steven L Puller, "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market," 2015. Working Paper.

Hu, Yingyao, "Microeconomic Models with Latent Variables: Applications of Measurement Error Models in Empirical Industrial Organization and Labor Economics," 2015. Working Paper.

Jochim, Ashley, Michael DeArmond, Betheny Gross, and Robin Lake, *How Parents Experience Public School Choice*, The Center on Reinventing Public Education, 2014.

Kadam, Sangram V. and Maciej H. Kotowski, "Multi-period Matching," 2015. Working Paper.

Kapor, Adam, "Distributional Effects of Race-Blind Affirmative Action," 2015. Working Paper.

Kennes, John, Daniel Monte, and Norovsambuu Tumennasan, "The Day Care Assignment: A Dynamic Matching Problem," *American Economic Journal: Microeconomics*, 2014, *6* (4), 362–406.

Ketcham, Jonathan D, Claudio Lucarelli, and Christopher A Powers, "Paying Attention or Paying Too Much in Medicare Part D," *American Economic Review*, 2015, *105* (1), 204–33.

_ , _ , Eugenio J Miravete, and M Christopher Roebuck, "Sinking, Swimming, or Learning to Swim in Medicare Part D," *American Economic Review*, 2012, *102* (6), 2639–2673.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel, "Comparison Friction: Experimental Evidence from Medicare Drug Plans," *Quarterly Journal of Economics*, 2012, *127* (1), 199–235.

Kurino, Morimitsu, "House Allocation with Overlapping Generations," *American Economic Journal: Microeconomics*, 2014, *6*, 258–89.

Leshno, Jacob D, "Dynamic Matching in Overloaded Waiting Lists," 2015. Working Paper.

Manski, Charles, *Identification for Prediction and Decision*, Harvard University Press, 2009.

_ and David A. Wise, *College Choice in America*, Harvard University Press, 1983.

Matzkin, Rosa L., "Nonparametric Identification," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics, Volume 6(B)*, Oxford: North Holland, 2007.

Nathanson, Lori, Sean Corcoran, and Christine Baker-Smith, "High School Choice in New York City: A Report on the School Choices and Placements of Low-achieving Students," 2013. Working Paper.

Neilson, Christopher, "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," 2013. Working Paper.

Pereyra, Juan Sebastian, "A Dynamic School Choice Model," *Games and Economic Behavior*, 2013, *80*, 100–114.

Robbins, Liz, "Lost in the School Choice Maze," *New York Times*, 2011.

Rossi, Peter E, Robert E McCulloch, and Greg M Allenby, "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 1996, *15* (4), 321–340.

Roth, Alvin E., "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica*, 2002, *70*, 1341–1378.

_ **and Marilda O. Sotomayor**, *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*, Cambridge: Econometric Society monographs, 1990.

Shum, Matthew, "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-cereals Market," *Journal of Economics & Management Strategy*, 2004, *13* (2), 241–272.

Spinnewijn, Johannes, "Heterogeneity, Demand for Insurance and Adverse Selection," 2015. Working Paper.

Sudhir, K and Nathan Yang, "Exploiting the Choice-Consumption Mismatch: A New Approach to Disentangle State Dependence and Heterogeneity," 2014. Working Paper.

Train, Kenneth, *Discrete Choice Methods with Simulation*, Cambridge University Press, 2009.

Ünver, M Utku, "Dynamic Kidney Exchange," *Review of Economic Studies*, 2010, *77* (1), 372–414.

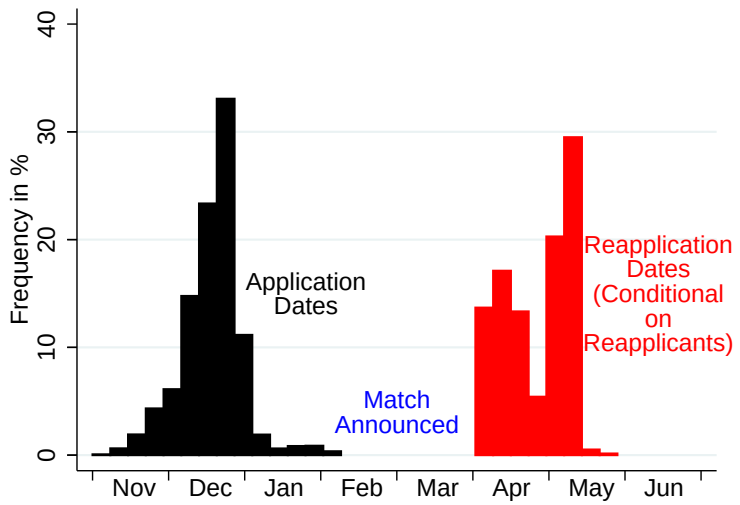
Varian, Hal, "Revealed Preference," in Aron A. Gottesman Michael Szenberg, Lall Ramrattan, ed., *Samuelsonian Economics and the Twenty-First Century*, Oxford: Oxford University Press, 2006.

Walters, Christopher, "The Demand for Effective Charter Schools," 2014. Working Paper.

Wiswall, Matthew and Basit Zafar, "Determinants of College Major Choice: Identification using an Information Experiment," *Review of Economic Studies*, 2015, *82* (2), 791–824.

Wong, Wing Hung, "Theory of Partial Likelihood," *The Annals of Statistics*, 1986, pp. 88–123.

Figure 1: Timeline of the First-round and Reapplication Process



Notes: In this figure, the left black histogram plots the distribution of dates at which applicants file initial applications. The right red histogram does the same for reapplication dates conditional on those who reapply. See Section 2.1 for discussions about this figure.

Table 1: Evolving School Choices

	Number	Fraction
All applicants in the first round	91,289	100%
Reapplicants in the aftermarket	6,430	7%
Reapplicants who exhibit preference reversals	4,564	5%
Reapplicants who exhibit surely nonstrategic preference reversals	3,464	4%
Lower bound on the fraction of nonstrategic preference reversals (4th row/3rd row)	76%	

Notes: This table shows how many applicants reapply, exhibit any choice reversals, and exhibit surely nonstrategic choice reversals between the first-round and the reapplication process. I say an applicant *exhibits choice reversals* if she reapplies against her initially assigned school s by ranking another school that is unranked or ranked below s in her initial application. An applicant *exhibits surely non-strategic choice reversals* if she exhibits any choice reversals and does not exhaust her preference list (rank 11 or fewer schools). As explained in the main text, surely non-strategic choice reversals can be rationalized only by real demand changes. See Section 2.1 for discussions about this figure.

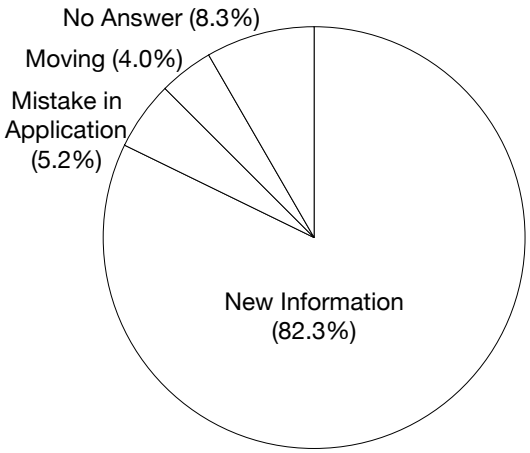
Table 2: Characteristics of Applicants and Reapplicants

	All applicants	All reapplicants	Reapplicants with preference reversals	Reapplicants with accepted reapplication
Average preference rank of initial assignment	4.0	5.7	4.3	5.5
Female	50%	55%	56%	53%
8th graders	95%	98%	98%	97%
Top 2% of grade 7 English, Language, Arts test	2.1%	0.6%	0.3%	1.0%
Grade 7 reading grade category				
High	15%	12%	10%	19%
Middle	72%	77%	78%	70%
Low	13%	11%	11%	11%
Living area				
Manhattan	12%	11%	11%	9%
Brooklyn	31%	34%	32%	32%
Queens	28%	31%	33%	40%
Bronx	22%	19%	20%	12%
Staten Island	7%	5%	3%	6%
Home language				
Chinese	10%	8%	8%	10%
Spanish	64%	62%	66%	56%
N	91,289	6,430	4,564	1,375
Ethnicity				
Asian	13%	14%	13%	17%
Hispanic	39%	41%	43%	39%
Black	34%	32%	33%	26%
White	14%	13%	10%	18%
N	83,047	5,914	4,269	1,283

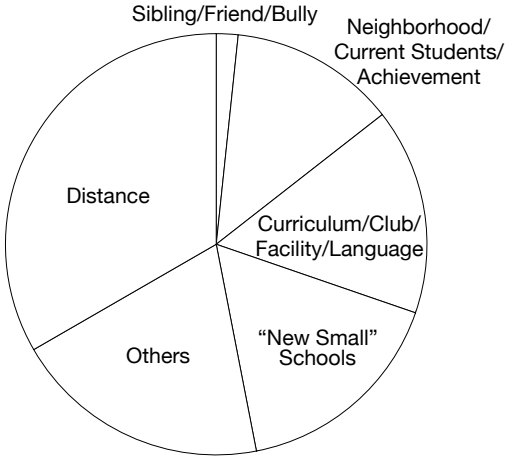
Notes: This table shows baseline characteristics of all applicants, applicants who reapply, applicants who reapply and exhibit choice reversals between their initial applications and reapplications, and applicants who reapply and are accepted. On "average preference rank of initial assignment," I assume the preference rank of being unassigned is 13, the worst possible rank plus one. Sample sizes in lower rows are smaller than that in upper rows because some characteristics are missing for some applicants. The definitions of test-score-related variables are in Appendix A.1. See Section 2.2 for discussions about this table.

Figure 2: Self-reported Reasons for Reapplication

(a) Self-reported Reasons for Reapplication



(b) Breakdown of "New Information"



Notes: Panel 2a classifies self-reported reasons for reapplying into main categories. Panel 2b focuses on the "New Information about Schools" category in Panel 2a and breaks it down into several sub-categories. "Distance" is different from "Moving" in that the former does not refer to any address change. The construction of the categories are explained in Appendix A.2. See Section 2.2 for discussions about this figure.

Table 3: Growing Response of Choices to Distance and Academic Achievement

		Initial applications		Reapplications	Change
		All applicants	Reapplicants	Reapplicants	Reapplicants
Distance (in miles)	1st choice school	4.5	4.9	3.9	-21.0%
	Schools ranked above initial assignment	5.1	5.1	4.0	-22.0%
	All ranked schools	4.8	5.3	4.0	-24.0%
Schools with low academic performance	1st choice school	12.8%	11.3%	10.5%	-6.8%
	Schools ranked above initial assignment	11.6%	10.1%	10.7%	6.4%
	All ranked schools	15.1%	14.1%	10.7%	-24.3%

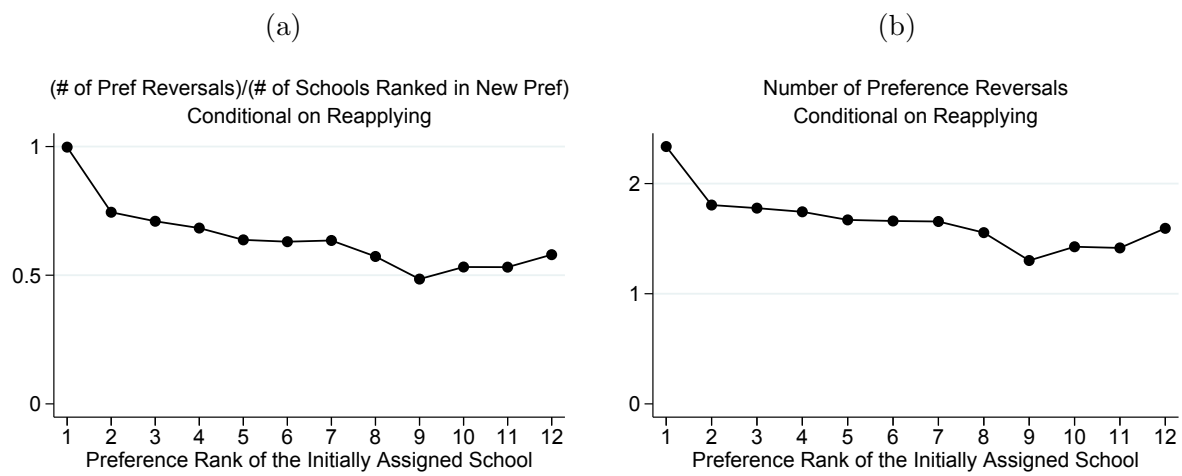
Notes: This table shows the average distance to and academic achievement level of ranked schools in initial applications and reapplications. *Schools with low academic performance* are "schools in need of improvement" in the official school brochure issued by NYC. Under the No Child Left Behind Act, New York State establishes annual performance goals in Mathematics and English Language Arts for all NYC public schools. Schools that do not meet these goals for two consecutive years are identified as schools in need of improvement. See Section 2.2 for discussions about this figure.

Table 4: Growing Response of Choices to School Characteristics

	Initial applications R2		Reapplications R2	Change
	All applicants	Reapplicants	Reapplicants	Reapplicants
Specification 1 (Location/borough dummies)	0.015	0.016	0.021	+31%
Specification 2 (1+Academic performance dummy)	0.027	0.038	0.054	+42%
Specification 3 (2+Program type dummies)	0.075	0.082	0.181	+121%
Specification 4 (3+Capacity dummies)	0.268	0.229	0.315	+38%
N (Schools)	750	750	750	

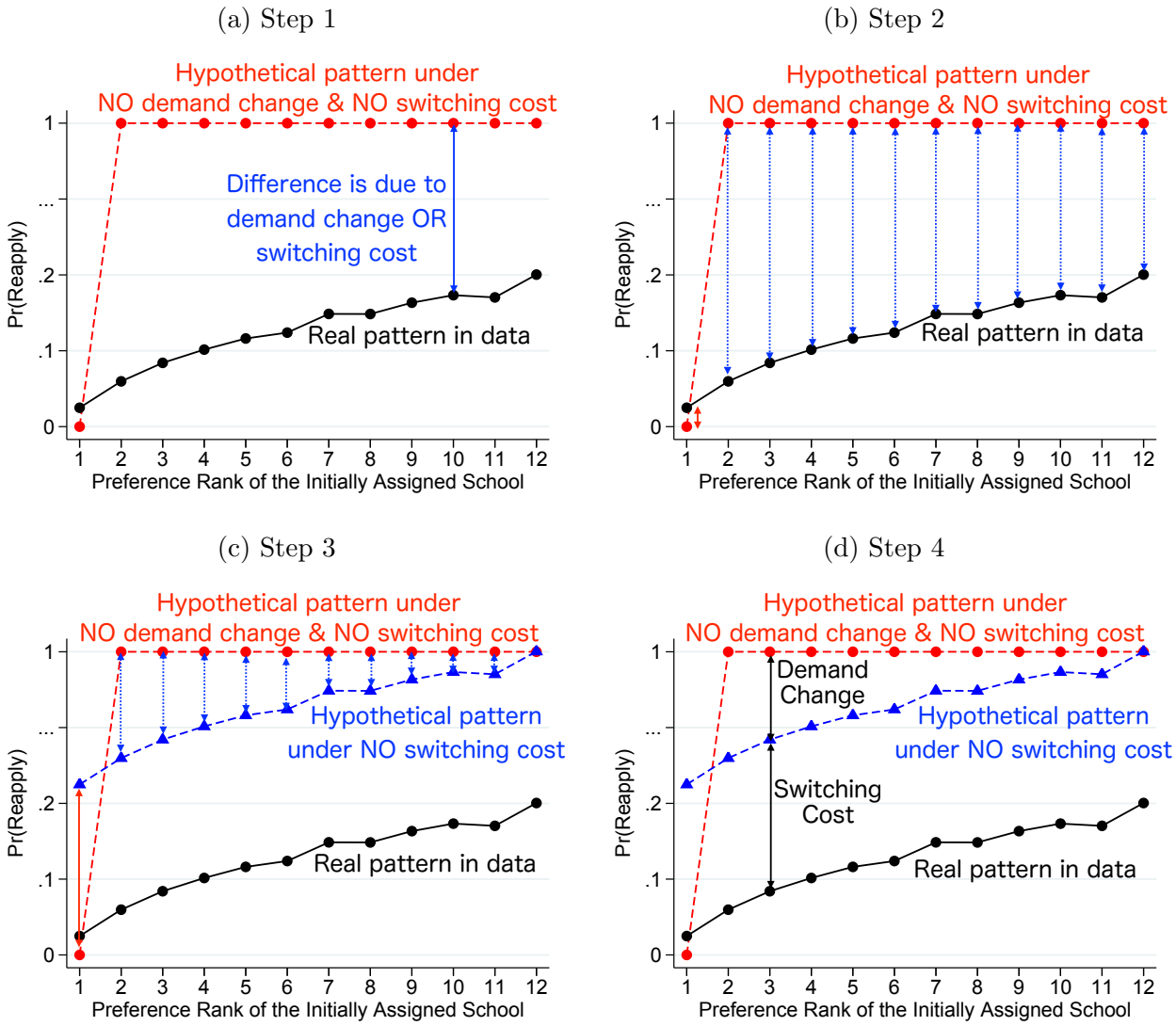
Notes: This table shows R^2's from school-level regressions of schools' first choice market shares on various sets of observable school characteristics. The *first choice market share* of a school is defined as the fraction of applicants who rank it first among all applicants who make a first choice. Rows correspond to different sets of school characteristics included in regressions. Columns correspond to different samples in different periods used to compute market shares. The details of included characteristics are explained in Appendix A.2. See Section 2.2 for discussions about this figure.

Figure 3: Falsification Check of the Empirical Model



Notes: This figure correlates two measures of the amount of choice reversals with the preference rank of the initially assigned school with respect to the initial preference (both are conditional on reapplicants). For each applicant who reapplies after being initially assigned to s, the number of *choice reversals* is defined as the number of schools t such that t is unranked or ranked below s in her initial application but ranked in her reapplication. See Section 3.1 for discussions about this figure.

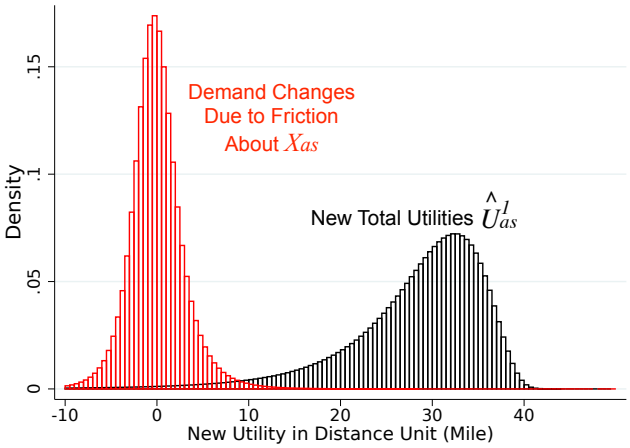
Figure 4: Identification and Evidence of Demand Changes and Switching Costs



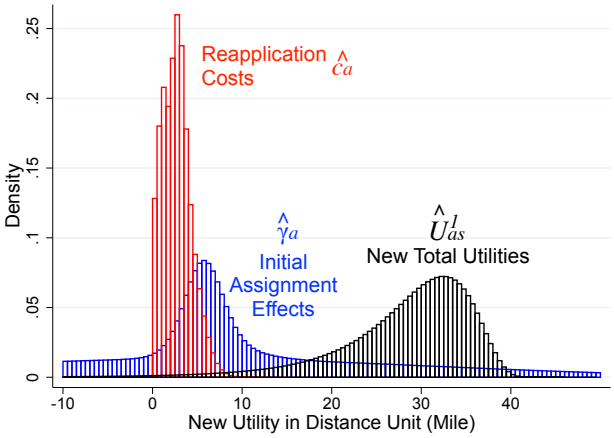
Notes: These figures explain how I separately identify demand changes and switching costs from the solid black line observed in the data. The solid black line correlates the conditional probability of reapplying to the preference rank of the initially assigned school with respect to the initial preference. See Section 3.2 for the identification argument using these figures.

Figure 5: Summary of Estimates

(a) Demand Changes due to Learning

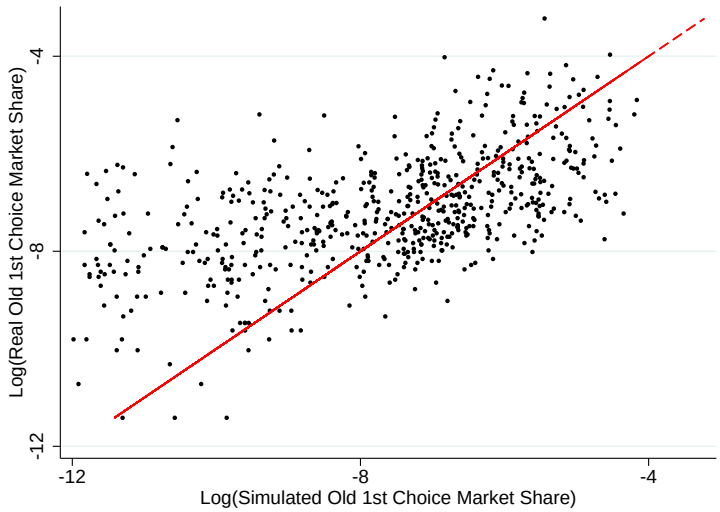


(b) Reapplication Costs & Initial Match Effects



Notes: Based on the rational expectation model in Section 3.1, Panel 5a plots the distributions of estimated overall new utilities ($\hat{U}^1_{as}$) and latent demand changes caused by frictions about observable school characteristics ($\Sigma_{k=1}^{K}\beta_{ak}f_{ak}X_{ask}$) for all (applicant a, school s) pairs. Panel 5b plots the distributions of estimated overall new utilities ($\hat{U}^1_{as}$), estimated reapplication costs ($\hat{c}_a$), and estimated initial assignment effects ($\hat{\gamma}_a$). Both panels are based on 50 simulations of the estimated model for each (applicant a, school s) pair. See Sections 3.1 and 3.3 for the details of the model and the estimation method, respectively. See Section 3.4 for discussions about this figure.

Figure 6: Fit (I) Initial Choices



Notes: This figure correlates the real first choice market share of each school in initial applications with the same share predicted by simulating the estimated model 50 times and averaging over them. The *first choice market share* of a school is defined as the fraction of applicants who rank it first among all applicants. See Section 3.4 for discussions about this figure.

Table 5: Fit (II) Reapplications, Choice Reversals, and Hidden Demand Changes

		Simulations of the estimated model			
		Model 1: Rational expectation		Model 2: Naive free expectation	
	Data	Estimated reapplication costs c_a & initial assignment effects Y_a	c_a=0 & estimated Y_a	Estimated c_a & estimated Y_a	c_a=0 & estimated Y_a
Reapplicants	7.0%	6.8% (0.15%)	30.8% (0.78%)	7.0% (0.46%)	39.6% (0.17%)
Reapplicants who exhibit choice reversals	5.0%	5.3% (0.18%)	23.4% (1.15%)	4.9% (1.55%)	25.7% (0.15%)
2nd row/1st row	71.4%	77.7%	76.1%	70.2%	65.0%

Notes: This table shows the fraction of applicants who reapply and exhibit choice reversals in the data and the estimated models with and without reapplication costs. I say an applicant *exhibits choice reversals* if she reapplies against her initially assigned school s by ranking another school that is unranked or ranked below s in her initial application. Each column for a model is based on simulating the corresponding estimated model 50 times and averaging over them. Simulation standard errors are in parentheses. See Sections 3.4 and 4.1 for discussions about this figure.

Table 6: Fit (III) Growing Correlation between Choices and School Characteristics

	Simulations of the estimated model			Data
	Initial app. R2	Reapp. R2	Change	Change
	Reapplicants	Reapplicants	Reapplicants	Reapplicants
Specification 1 (Location/borough dummies)	0.018	0.020	+11.1%	+31.3%
Specification 2 (1+Academic performance dummy)	0.039	0.041	+5.1%	+42.1%
Specification 3 (2+Program type dummies)	0.096	0.250	+160.4%	+120.7%
Specification 4 (3+Capacity dummies)	0.587	0.520	-11.4%	+37.6%
N (schools)	750	750		

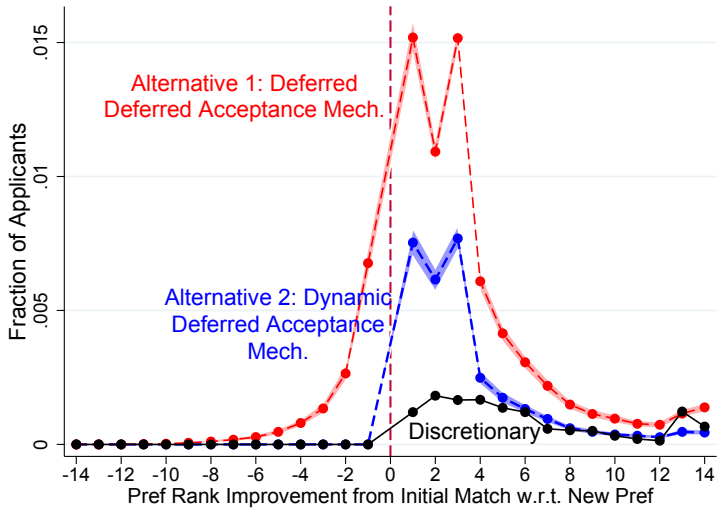
Notes: Based on the rational expectation model, this table shows changes in R^2's from school-level regressions of schools' first choice market shares on observable school characteristics. The *first choice market share* of a school is defined as the fraction of applicants who rank it first among all applicants who make a first choice. Rows correspond to different sets of school characteristics included in regressions. Columns correspond to different periods used to compute market shares. For the model, I compute predicted shares by simulating the estimated model 50 times and averaging over them. See Section 3.4 for discussions about this figure.

Table 7: Welfare Costs of Ignoring Demand Changes

Real 1st-round assignment vs frictionless 1st-round assignment under new demand					
Model 1: Rational expectation			Model 2: Naive free expectation		
% of winners	% of losers	Avg util change	% of winners	% of losers	Avg util change
47.5% (0.14%)	11.1% (0.11%)	+1.60 miles (0.01 miles) +0.29 utility SD (0.08 utility SD)	48.9% (0.22%)	12.2% (0.13%)	+1.00 miles (0.01 miles) +0.16 utility SD (0.001 utility SD)

Notes: This table shows welfare changes from the real first-round assignment based on old demand to the counterfactual "frictionless benchmark" that is defined as what would have been produced by the same first-round assignment mechanism had families made choices based on their new demand after learning. "Winners" ("losers") are defined as applicants who be better (worse, respectively) off under the frictionless benchmark, compared with the real first-round assignment with respect to new demand U^1_{as}. Average utility changes are also measured by new demand U^1_{as}, which is assumed to be quasi-linear in the distance between the family and the school locations. "Utility SD" is measured by the standard deviation in the distribution of utilities from all schools for each applicant. Simulation standard errors over 50 simulations are in parentheses. See Section 4.1 for discussions about this figure.

Figure 7: Centralized vs Discretionary Reapplication Processes (I)



Notes: In this figure, each line plots the distribution of improvements of the preference rank of the finally assigned school under each of the three ways to accommodate observed choice changes: the real discretionary reapplication process, the dynamic deferred acceptance mechanism, and the deferred deferred acceptance mechanism. Section 4.2 defines the latter two mechanisms. The common status quo for these comparisons is the initial assignment in the initial application process via the static deferred acceptance algorithm. The preference rank in April is defined with respect to new preference $\succ_a^1$ defined in Section 4.2. This distribution is conditional on applicants who get different assignments under the two mechanisms. The shaded area around a line indicates the 95% simulation confident interval over 200 simulations of lottery numbers used by the mechanisms to break ties in priorities. See Section 3.2 for details of the use of lottery numbers in the NYC system. There is no shaded area around the black line for the real reapplication process since I observed only one realization of the real reapplication process and it is not possible to simulate it. See Section 4.2 for discussions about this figure.

Table 8: Centralized vs Discretionary Reapplication Processes (II)

		Model 1: Rational expectation			Model 2: Naive free expectation			Descriptive counterfactual	
		% of winners	% of losers	Avg util change	% of winners	% of losers	Avg util change	% of winners	% of losers
Real discretionary aftermarket				N/A				1%	0%
Dynamic DA mech.	With estimated reapplication costs	3.0% (0.1%)	0% (0%)	+0.15 miles (0.01 miles) +0.02 utility SD (0.003 utility SD)	7.0% (5.6%)	0% (0%)	+0.20 miles (0.14 miles) +0.031 utility SD (0.021 utility SD)	4%	0%
Deferred DA mech.	With estimated reapplication costs	5.4% (0.1%)	3.4% (0.04%)	+0.11 miles (0.003 miles) +0.01 utility SD (0.002 utility SD)	7.7% (2.7%)	3.6% (4.2%)	+0.23 miles (0.21 miles) +0.032 utility SD (0.009 utility SD)	8%	2%

Notes: This table summarizes the effects of discretionary and centralized reapplication processes on family welfare in the form of the fractions of winners and losers, who would be better and worse off, respectively, by the introduction of each reapplication process compared with the initial match. The table also reports the average utility change in distance-equivalent utility units. I define winners, losers, and utility changes in terms of new utilities U^1_{as} after learning (including initial assignment effects γ_a). "Utility SD" is measured by the standard deviation in the distribution of utilities from all schools for each applicant. Each row for a reapplication process is based on simulating the estimated model 50 times under the reapplication process and averaging over the simulations. Simulation standard errors are in parentheses. Simulation standard errors for the descriptive evaluation are negligible, as suggested by simulation standard errors in Figure 7. The last two columns are based on the descriptive counterfactual analysis detailed in Figure 7. See Section 4.2 for discussions about this table.

Table 9: Dynamic Reapplication Processes and Demand-side Inertia

		Model 1: Rational expectation			Model 2: Naive free expectation		
		% of winners	% of losers	Avg util change	% of winners	% of losers	Avg util change
Dynamic DA mech.	With estimated reapplication costs	3.0% (0.1%)	0% (0%)	+0.15 miles (0.01 miles) +0.02 utility SD (0.003 utility SD)	7.0% (5.6%)	0% (0%)	+0.20 miles (0.14 miles) +0.031 utility SD (0.021 utility SD)
	With no reapp. costs (Infeasible)	17.3% (0.2%)	0% (0%)	+0.68 miles (0.08 miles) +0.11 utility SD (0.01 utility SD)	25.2% (0.2%)	0% (0%)	+0.71 miles (0.01 miles) +0.11 utility SD (0.001 utility SD)
	Dilution by reapp. costs	-83%	N/A	-78%	-72%	N/A	-72%
Deferred DA mech.	With estimated reapplication costs	5.4% (0.1%)	3.4% (0.04%)	+0.11 miles (0.003 miles) +0.01 utility SD (0.002 utility SD)	7.7% (2.7%)	3.6% (4.2%)	+0.23 miles (0.21 miles) +0.032 utility SD (0.009 utility SD)
	With no reapp. costs (Infeasible)	47.5% (0.1%)	11.1% (0.1%)	+1.60 miles (0.01 miles) +0.29 utility SD (0.08 utility SD)	48.9% (0.22%)	12.2% (0.13%)	+1.00 miles (0.01 miles) +0.16 utility SD (0.001 utility SD)
	Dilution by reapp. costs	-89%	-69%	-93%	-84%	-70%	-77%

Notes: This table summarizes the effects of discretionary and centralized reapplication processes on family welfare in the form of the fractions of winners and losers, who would be better and worse off, respectively, by the introduction of each reapplication process compared with the initial match. The table also reports the average utility change in distance-equivalent utility units. "With no reapplication costs" means making reapplication costs c_a zero (while keeping initial assignment effects γ_a at the estimated value). I define winners, losers, and utility changes in terms of new utilities U^1_{as} after learning (including initial assignment effects γ_a). "Utility SD" is measured by the standard deviation in the distribution of utilities from all schools for each applicant. Each row for a reapplication process is based on simulating the estimated model 50 times under the reapplication process and averaging over the simulations. Simulation standard errors are in parentheses. See Section 4.2 for discussions about this table.

A Appendix

A.1 Data

Applications, reapplications, and assignments. The data come from the NYC Department of Education (DOE) office. It contains information on students and their characteristics, schools/programs and their characteristics, the initial rank-order preference lists submitted by the students, students' priority statuses at schools, the initial assignment of students to schools, whether each student reapplies, the new rank-order preference lists submitted by the students who reapply, and the dates of initial applications and reapplications. For applicants who reapply, the data also contain written reapplication reasons as self-reported by reapplicants.

Applicant characteristics. The records from the DOE contain the street address, previous and current grade, gender, ethnicity, whether the student was enrolled in a public middle school, scores in middle school standardized tests, limited English proficiency status, and special education status. In addition, applicants are categorized into one of three categories based on their score on the seventh grade standardized reading test: top 16 percent (high), middle 68 percent (middle), and bottom 16 percent (low). This categorization is made mainly for admissions at "educational option" programs, which I explain below. Applicants are also categorized according to whether they are in the top 2% of the grade 7 English Language Arts (ELA) test.

School/program characteristics. In the NYC high school system, there are three types of schools: schools that actively evaluate applicants and submit a ranking to the mechanism; schools that do not evaluate applicants, and instead order students by priorities, which are determined not at the school, but by the DOE; and schools at which a fraction of seats are reserved for students who are explicitly ranked by the school, while the rest are automatically categorized into priority groupings set by the DOE. "Screened" and "audition" schools are examples of the first type of school, at which staff review applicants based on criteria such as seventh grade academic performance, attendance, disciplinary actions, auditions, portfolio submissions, and interviews. "Unscreened" schools are examples of the second type of school. Priorities are based on geographic location, current middle school, or other criteria. Finally, the third class of schools, "educational option," are permitted to rank students for half of their positions, and are required to admit students according to priorities for the other half. Nearly half of all schools are educational option, and more than half of total district capacity is at schools that do not actively rank students. When priorities are used at unscreened and educational option programs, many students fall into the same priority class. Ties or indifferences within each priority group are broken by a single lottery commonly used by all

schools.

"Low academic performance" is a dummy for being categorized as one of the "schools in need of improvement" in the official school brochure issued by the DOE. Under the No Child Left Behind Act, New York State establishes annual performance goals in Mathematics and English Language Arts for all NYC public schools. Schools that do not meet these goals for two consecutive years are identified as schools in need of improvement.

I say a school program is "new" if it was created in or after 2002. Otherwise, I call it "old." Program capacities are not provided in the data. I have estimated program capacities from the assignment. Specifically, I define the capacity of a program as the number of students assigned to it in the main application process; this method should be justified because most schools reject at least some applicants in the application process and the capacity of any of these schools should be the same as the number of applicants assigned to it. I say a school is "tiny" if it is in the bottom 25% of the distribution of school capacities among schools in the sample. "Small," "medium," and "large" correspond to 25-50%, 50-75%, and 75-100%respectively.

Distance. ArcGIS (with the address-set in the Business Analyst toolbox version 10.0) is used to geocode student and school addresses and calculate the distance between them on the road network. An exact match was first used to determine if a student's address can be geocoded precisely. If the results were unreliable, the student is assigned to the centroid of the zip-code. The vast majority of students were placed at the roof-top level. The OD Cost matrix tool in the Network Analyst toolbox was used to compute the distance by road for each student-school pair. The road network is also obtained from the Business Analyst toolbox.

A.2 A First Look at Evolving Choices: Details

A.2.1 Construction of Categories in Figure 2

For creating Figure 2, I randomly picked 10% of reapplication reasons and grouped them into categories mentioned in the Figure. To deal with multiple reasons reported by a single applicant, I first attached an equal weight of one to each reapplicant. When a single reapplicant refers to multiple reasons, I divide the reapplicant's weight equally across all referred reasons. "Moving" contains all reasons that refer to address changes after the initial application process. "Distance" includes the other distance-related reasons that do not refer to address changes. "Mistake in Application" contains reasons that mention mistakes such as misspelling program codes in the initial application or not intending to apply for the initially assigned school. "Current students" contains reasons related to current students at

the initially assigned school. For example, some families complain that current students are so scary that they do not want to go to the initially assigned school.

A.2.2 Construction of Regressors in Table 4

In Table 4, "Location (borough) dummies" are full dummies for being in each of Manhattan, Brooklyn, Queens, Bronx, and Staten Island. "Academic performance dummy" is the low academic performance dummy defined in Appendix A.1. "Program type dummies" are full dummies for program types explained in Appendix A.1. "Capacity dummies" are dummies for "tiny", "small", and "medium" defined in Appendix A.1.

A.3 Recovering Evolving Demand: Details

A.3.1 Identification: Details

Section 3.2 explains how to use admissions lotteries to separately identify switching costs and demand changes. Figure 4 converts the identification logic into suggestive descriptive evidence of significant demand changes and switching costs, but only under the simplifying assumption that initially assigned schools are completely randomly assigned. In the NYC school choice system, however, initial assignments are not purely random since initial assignments also depend on non-random preferences of applicants and their priorities at schools. This section extends the evidence of switching costs and demand changes in Figure 4 to the more complicated real case with partially random initial assignments.

As already explained in Section 3.2, as far as identification is concerned, partial randomization does not cause any serious problem: I can repeat the same analysis as in Figure 4 *conditional on each applicant type* (i.e., entire preference list and priorities at all schools) to separately identify switching costs and demand changes for that type. There is a problem, however, when I try to extend the empirical implementation of the identification argument in Figure 4 to partially random initial assignments. Since applicant type is a high-dimensional object (e.g. with 750 schools, the number of possible preferences is $750 \times 749 \times 748 \times ...$), few applicants usually share the same type in any reasonably sized data. This fact makes it infeasible to draw a conditional-on-full-type version of Figure 4. I therefore need to resort to a different strategy that conditions on something coarser to make initial assignments random.

Following the standard notation in econometrics of program evaluation, let $D_{as} = 1$ if applicant a is assigned school s; $D_{as} = 0$ otherwise. $\succ_a^0$ denotes applicant a's old preference a submits in her initial application. $s \succ_a^0 s'$ means applicant a prefers school s over s' in

her initial application. Let ρ_{as} be a's priority at s (before tie-breaking by lotteries). Define $First_s \equiv \{a|s \succ_a^0 s'$ for all school $s' \neq s$ and there exists applicant a' such that $\rho_{as} = \rho_{a's}$ and $D_{as} \neq D_{a's}\}$ as the set of applicants who rank school s first and are in s's "marginal priority group" where some applicants are assigned s but others are not though all of them share the same priority at s. Since all applicants in $First_s$ rank school s first and share the same priority at s, whether they get assigned to s should be determined solely by their lottery numbers. Therefore, assignments to school s within $First_s$ can be thought as if being randomly assigned in a randomized controlled trial.

I use this strategy to create a structural or causal version of Figure 4 as follows. Let me pool $First_s$ across all schools into a single sample $\cup_s First_s$. Within $\cup_s First_s$, consider the following regression (linear probability model) or its nonlinear logit or probit version:

$$Y_a = \beta D_a + \Sigma_s \alpha_s X_{as} + \epsilon_a,$$

where $Y_a \equiv 1\{$applicant a reapplies$\}$, $D_a \equiv 1\{$applicant a is assigned to her first choice school$\}$, and $X_{as} \equiv 1\{$applicant a ranks school s first$\}$. By the above argument, D_a is asymptotically randomly assigned conditional on X_{as}'s (the identity of the first choice school) and thus estimated $\hat{\beta}$ from the above regression is causally or structurally interpretable. In fact, Table A.3 confirms that conditional on being in $\cup_s First_s$ and X_{as}'s (the identity of the first choice school), baseline covariates are balanced between applicants who do and do not get first choice offers. Such covariate balance is lost, however, if I run the same regression with no controls. This suggests that D_a is indeed conditionally randomly assigned as intended.

Table A.4 reports $\hat{\beta}$ alongside the corresponding marginal effect estimates from the probit and logit versions of the above regression. The causal effect of being assigned to the first choice school on the probability of reapplying is precisely estimated and about 6%. The omitted variable bias appears to be small and the descriptive effect of being assigned to the first choice school on the probability of reapplying is about 7%.

Finally, let $\hat{\alpha} \equiv \dfrac{\Sigma_s |First_s| \hat{\alpha}_s}{\Sigma_s |First_s|}$ be the weighted average of $\hat{\alpha}_s$ across all schools with the weight being the number of applicants in $First_s$, who are randomly assigned to or rejected by each school. Each $\hat{\alpha}_s$ is the conditional probability of reapplying conditional on being assigned to a non-first-choice school and in $First_s$. Recall that $\hat{\beta}$ is a weighted average of the causal effects of being assigned to the first choice school on the probability of reapplying. Therefore, $\hat{\alpha} + \hat{\beta}$ and $\hat{\alpha}$ are (weighted averages of) the conditional probabilities of reapplying conditional on being assigned to the first choice school and a non-first-choice school, respectively, and in $\cup_s First_s$, where applicants are randomly assigned between the first choice school and a non-first-choice school.

Figure A.6 plots $\hat{\alpha}+\hat{\beta}$ and $\hat{\alpha}$ at $x = 1$ and $x = lower$, respectively, focusing on applicants in $\cup_s First_s$. Since applicants in the sample $\cup_s First_s$ used for creating Figure A.6 are purely randomly assigned between the two points on the x axis, Figure A.6 can be interpreted as a causal or structural version of Figure 4. By the logic in Figure 4 and Section 3.2, Figure A.6 structurally confirms that there are both significant switching costs and demand changes (at least for the subpopulation of applicants I focus on).

More formally, I can test the presence of switching costs and demand changes. If there were neither switching costs nor demand changes, all applicants assigned to a lower choice school would reapply while none of applicants assigned to the first choice school would reapply. This behavioral hypothesis corresponds to the statistical hypothesis that $\alpha+\beta = 0$ and $\alpha = 1$. Since α and β are linear combinations of linear regressions coefficients α_s's and β_s's, null hypotheses $H_0 : \alpha + \beta = 0$ and $H_0 : \alpha = 1$ are linear hypotheses. I can test these hypotheses by the usual Wald test.

Similarly, if there were no switching costs (but there may be demand changes), conditional on being in $\cup_s First_s$,

$$
\begin{aligned}
&\alpha + \beta \\
&= \Pr(a \text{ reapplies}|s_a^0 \text{ is } a\text{'s old 1st choice}) \\
&= \Sigma_{K\geq 2} \Pr(a \text{ does not reapply}|\ s_a^0 \text{ is } a\text{'s old } K\text{-th choice}) \\
&\quad \text{(by the same argument as in Section 3.2)} \\
&\geq \Sigma_{K\geq 2} \Pr(a \text{ does not reapply}|\ s_a^0 \text{ is } a\text{'s old } K\text{-th choice}) \\
&\quad \times \Pr(s_a^0 \text{ is } a\text{'s old } K\text{-th choice}|\ s_a^0 \text{ is } \textit{not } a\text{'s old 1st choice}) \\
&= 1 - \alpha,
\end{aligned}
$$

where the inequality is because $\Pr(a \text{ does not reapply}|\ s_a^0 \text{ is } a\text{'s old } K\text{-th choice}) \geq 0$ and $\Pr(s_a^0 \text{ is } a\text{'s old } K\text{-th choice}|\ s_a^0 \text{ is } \textit{not } a\text{'s old 1st choice}) \leq 1$. Thus, the behavioral hypothesis of no switching costs corresponds to the statistical hypothesis that $\alpha + \beta \geq 1 - \alpha$.

Table A.5 shows that the Wald test rejects both the null hypothesis that there are no switching costs ($\alpha + \beta = 1 - \alpha$) and the null hypothesis that there are neither switching costs nor demand changes ($\alpha + \beta = 0$ or $\alpha = 1$).

A.3.2 Estimation: Details

Likelihood Derivation and Estimation Procedure

In this section, I derive partly analytical likelihood functions from my empirical models (in Section 3.1) combined with the distributional assumptions in Section 3.3. The likelihood

functions at the family level are computed for a sequence of choices from the initial application period $t = 0$ to the reapplication period $t = 1$, since due to initial assignment effects and reapplication costs, the likelihood of a choice in $t = 1$ depends on the choice and initial assignment in the previous period $t = 0$. Let s^t_{al} be applicant a's l-th choice school in her period t preference $\succ^t_a$ $(t = 0, 1)$ and let $\#^t_a$ be the number of schools a ranks in $\succ^t_a$.

In period 0 (the initial application period), every applicant a in the sample submits $\succ^0_a$ and it implies that $U^0_{as^0_{al}} > U^0_{as}$ for every $l = 1, ..., \#^0_a$ and every s with $s^0_{al} \succ^0_a s$, including unranked schools. By the formula for logit choice probabilities (Train 2009 chapter 3), conditional on $(\mathbf{U}^0 \equiv (U^0_s)_s, (\beta_{ak}(1 + f_{ak}))_{k=1,...,K})$, the likelihood of observing $\succ^0_a$ is

$$\mathcal{L}^0_a \equiv \Pi^{\#^0_a}_{l=1} \frac{\exp(U^0_{s^0_{al}} + \Sigma^K_{k=1}\beta_{ak}(1 + f_{ak})X_{as^0_{al}k})}{\Sigma_{s \text{ with } s^0_{al} \succsim^0_a s} \exp(U^0_s + \Sigma^K_{k=1}\beta_{ak}(1 + f_{ak})X_{ask})}.$$

In period 1 (the reapplication process), even if applicant a does not reapply and does not submit new preference $\succ^1_a$, she provides useful information. In particular, in the rational expectation model, applicant a reapplies if equation (3) in the main text holds, implying that conditional on $(\mathbf{U}^1 \equiv (U^1_s)_s, (\beta_{ak}(1 + f_{ak}))_k, (\beta_{ak}f_{ak})_k, \boldsymbol{\epsilon}^0_a \equiv (\epsilon^0_{as})_s, c_a/p_a, \gamma_a)$, the likelihood of observing a's non-reapplication is

$$\mathcal{L}^{1,no\ reapp,rational}_a \equiv 1 - \int \mathbb{1}\{\text{equation (3) holds}\} dF_\theta,$$

where F_θ is the distribution of utility function arguments, parametrized by θ. If applicant a reapplies and submits $\succ^1_a$, then it implies that (I) equation (3) holds, (II) $p_{as^1_{al}} U^1_{as^1_{al}} > p_{as} U^1_{as}$ for each $l = 1, ..., \#^1_a$ and every unranked $s \neq s^0_a$, and (III) if a does not exhaust her reapplication preference, i.e., $\#^1_a < 3$ (recall any reapplicant can rank up to three schools in her reapplication preference), then $U^1_{as^0_a} + \gamma_a > U^1_{as}$ for all $s(\neq s^0_a)$ which a does not rank in $\succ^1_a$. Thus, the likelihood of observing a's reapplication and $\succ^1_a$ is

$$\begin{aligned}
&\mathcal{L}^{1,reapp,rational}_a \\
&\equiv \underbrace{\int \mathbb{1}\{\text{equation (3) holds}\} \times \Pi^{\#^1_a}_{l=1} \mathbb{1}\{p_{as^1_{al}} U^1_{as^1_{al}} > p_{as} U^1_{as} \text{ for all unranked } s \neq s^0_a\} dF'_\theta}_{\text{likelihood of (I) and (II) conditional on (III)}} \\
&\times \underbrace{[\mathbb{1}\{\#^1_a < 3\} \frac{\exp(U^1_{s^0_a} + \Sigma^K_{k=1}\beta_{ak}X_{as^0_a k} + \gamma_a + \epsilon^0_{as^0_a})}{\Sigma_{s \text{ unranked in } \succ^1_a} \exp(U^1_s + \Sigma^K_{k=1}\beta_{ak}X_{ask} + \mathbb{1}\{s = s^0_a\}\gamma_a + \epsilon^0_{as})} + (1 - \mathbb{1}\{\#^1_a < 3\})]}_{\text{likelihood of (III)}},
\end{aligned}$$

where F'_θ is the distribution of relevant utility function arguments *conditional on (III)*. If a exhausts her reapplication preference, i.e., $\#^1_a = 3$, then F'_θ is the same as the unconditional

distribution.

In the naive free expectation model, not reapplying implies that $U^1_{as^0_a} + c_a/p_a + \gamma_a > U^1_{as}$ for every $s \neq s^0_a$. By the formula for logit choice probabilities, the likelihood of observing a's non-reapplication is

$$\mathcal{L}^{1,no\ reapp,naive}_a \equiv \frac{\exp(U^1_{s^0_a} + \Sigma^K_{k=1}\beta_{ak}X_{as^0_a k} + \epsilon^0_{as^0_a} + c_a/p_a + \gamma_a)}{\Sigma_s \exp(U^1_s + \Sigma^K_{k=1}\beta_{ak}X_{ask} + \mathbb{1}\{s = s^0_a\}(c_a/p_a + \gamma_a) + \epsilon^0_{as})}.$$

If applicant a reapplies and submits $\succ^1_a$, then it implies that (i) it is *not* the case that $U^1_{as^0_a} + c_a/p_a + \gamma_a > U^1_{as}$ for every $s \neq s^0_a$, (ii) $U^1_{as^1_{al}} > U^1_{as}$ for each $l = 1, ..., \#^1_a$ and every s with $s^1_{al} \succ^1_a s$, and (iii) if a does not exhaust her reapplication preference, i.e., $\#^1_a < 3$ (recall any reapplicant can rank up to three schools in her reapplication preference), then $U^1_{as^0_a} + \gamma_a > U^1_{as}$ for all $s(\neq s^0_a)$ which a does not rank in $\succ^1_a$. Conditional on $(\mathbf{U}^1, (\beta_{ak}(1+f_{ak}))_k, (\beta_{ak}f_{ak})_k, \boldsymbol{\epsilon}^0_a, c_a/p_a, \gamma_a)$, the likelihood of observing a's reapplication and $\succ^1_a$ is

$$\begin{aligned}
&\mathcal{L}^{1,reapp,naive}_a \\
&\equiv \int (1 - \mathbb{1}\{U^1_{as^0_a} + c_a/p_a + \gamma_a > U^1_{as} \text{ for every } s \neq s^0_a\}) \\
&\quad \underbrace{\times \Pi^{\#^1_a}_{l=1} \mathbb{1}\{U^1_{as^1_{al}} > U^1_{as} \text{ for every } s \text{ with } s^1_{al} \succ^1_a s\} dF'_\theta}_{\text{likelihood of (i) and (ii) conditional on (iii)}} \\
&\quad \times \underbrace{[\mathbb{1}\{\#^1_a < 3\} \frac{\exp(U^1_{s^0_a} + \Sigma^K_{k=1}\beta_{ak}X_{as^0_a k} + \gamma_a + \epsilon^0_{as^0_a})}{\Sigma_{s \text{ unranked in } \succ^1_a} \exp(U^1_s + \Sigma^K_{k=1}\beta_{ak}X_{ask} + \mathbb{1}\{s = s^0_a\}\gamma_a + \epsilon^0_{as})} + (1 - \mathbb{1}\{\#^1_a < 3\})]}_{\text{likelihood of (iii)}}.
\end{aligned}$$

For each model, integrating over $(\mathbf{U}^0, \mathbf{U}^1, (\beta_{ak}(1+f_{ak}))_k, (\beta_{ak}f_{ak})_k, \boldsymbol{\epsilon}^t_a, c_a/p_a, \gamma_a)$ and all applicants gives the full, unconditional likelihood as follows:

$$\mathcal{L}(\theta) \equiv \Pi_a \int \mathcal{L}^0_a [\mathbb{1}\{a \text{ reapplies}\}\mathcal{L}^{1,reapp}_a + (1 - \mathbb{1}\{a \text{ reapplies}\})\mathcal{L}^{1,no\ reapp}_a] dF_\theta.$$

For estimating θ under each model, I find $\hat{\theta}$ that maximizes the simulated version of the logarithm of the likelihood function. More precisely, for easing the computational burden, I take a two step procedure. The first step estimates the part of θ related to $\mathbf{U}^0$ and $\beta_{ak}(1+f_{ak})$ by maximizing the simulated version of the logarithm of the partial likelihood $\Pi_a \int \mathcal{L}^0_a dF_\theta$. Taking the estimated part of θ as given, the second step estimates the remaining part of θ by maximizing the simulated version of the logarithm of the remaining conditional likelihood for period 1, i.e., $\Pi_a \int \mathbb{1}\{a \text{ reapplies}\}\mathcal{L}^{1,reapp}_a + (1 - \mathbb{1}\{a \text{ reapplies}\})\mathcal{L}^{1,no\ reapp}_a dF_\theta$. This two step approach is legitimate since (1) the estimation target of the first step ($\mathbf{U}^0$

and $\beta_{ak}(1 + f_{ak})$) is identified by period 0 preference $\succ_a^0$ alone, and maximizing the period 0 partial likelihood gives us consistent estimates with conservative standard errors, and (2) the estimation target of the second step depends only on the period 1 conditional likelihood, and standard results for estimators based on maximizing a partial or conditional likelihood guarantee its consistency and asymptotic normality (Wong, 1986).

Finally, when simulating old utilities U_{as}^0's in the second step, I do so conditional on applicant a's observed initial application preference $\succ_a^0$.[36] This conditional simulation seems more appropriate than unconditional simulation for incorporating the fact that observed initial assignment s_a^0 (which is used in the estimation) is one of the twelve most preferred schools in initial application preference $\succ_a^0$ and underlying old utilities U_{as}^0's. If I simulate old utilities unconditionally, observed initial assignment s_a^0 is often not such a preferred school in simulated period 0 utilities (especially because the number of schools is large). I then need large reapplication costs and initial assignment effects to explain the low reapplication rate, which may result in overestimating reapplication costs and initial assignment effects.

Estimation Sample Construction

Schools. 755 school programs are ranked by some students in the initial application process. Among them, I had to drop 5 schools that I could not geocode because because I did not receive their address information. I also dropped applicants' choices associated with these dropped schools. This left 750 schools in the estimation and counterfactual sample.

Applicants. 91289 students make applications and submit preferences in the initial application process. Starting from them, I needed to focus on students who rank at least one in-sample school in the initial application process, since other students reveal no information about their preferences over in-sample schools. I excluded students in the top 2 percent of the grade 7 English language arts test score by the following reason: any students in the top 2 percent are guaranteed assignment to their first choice if they rank a program of a particular type, known as "educational option," as their first choice. A student who does not prefer an educational option program as her top choice may thus have an incentive to strategize and rank it as her top choice so that she receives it. If so, their stated preferences do not follow the truth-telling assumption necessary for identification. I also had to drop 9th graders, who participate in the mechanism for the second time and are potentially subject to different types of learning or effects of a prior assignment on utilities, and students from private schools, for whom I do not observe necessary demographic information.

[36] In particular, within each simulation, I first simulate $\beta_{ak}(1+f_{ak})$ and U_s^0 unconditionally and take them as fixed. Then I simulate unobserved utility shocks ϵ_{as}^0's conditional both on initial application preference $\succ_a^0$ and the simulated values of $\beta_{ak}(1 + f_{ak})$ and U_s^0 so that the resulting period 0 utilities U_{as}^0's become distributed conditional on initial application preference $\succ_a^0$.

Finally, I dropped students for whom no address information was available, because I could not geocode them. As a result of this process, 76250 students remain in the estimation and counterfactual sample.

Choices. A tiny number of choices are recorded without preference ranks. I had to drop these choices.

A.4 Theory for the Counterfactual Analysis

A.4.1 Model

There are a finite set A of applicants and a finite set S of schools. Each school $s \in S$ has a strict preference/priority relation $\succ_s$ over the set of subsets of A. In many settings including the NYC system, school priorities are coarse or weak and lotteries are used to break ties or indifferences. In the theoretical analysis, I do not explicitly consider the tie-breaking process and take ex post strict priorities as given.

I assume that the preference relation of each school is *responsive with capacity* q_s (Roth and Sotomayor, 1990), i.e.,

(1) For any $a, \bar{a} \in A$, if $\{a\} \succ_s \{\bar{a}\}$, then for any $A' \subseteq A \setminus \{a, \bar{a}\}$, $A' \cup \{a\} \succ_s A' \cup \{\bar{a}\}$,

(2) For any $a \in A$, if $\{a\} \succ_s \varnothing$, then for any $A' \subseteq A$ such that $|A'| < q_s$, $A' \cup \{a\} \succ_s A'$, and

(3) $\varnothing \succ_s A'$ for any $A' \subseteq A$ with $|A'| > q_s$.

If a school's preferences are responsive, then that school acts as if it has preferences over individual applicants and a quantity constraint, and the school takes the available highest-ranking applicants up to that quantity constraint. In addition, I assume that every applicant is acceptable to every school because I am primarily interested in the assignment of applicants to public schools as in the NYC system. The preference profile of all schools is denoted by $\succ_S \equiv (\succ_s)_{s \in S}$. In this model, schools' preferences can be interpreted as their intrinsic preferences or priorities exogenously given by law or the authorities. In NYC, the former interpretation is appropriate for some schools (e.g., "screened" schools) while the latter interpretation is suitable for other schools (e.g., "unscreened" schools). The model and the results are applicable to both interpretations.

On top of the above usual structure, the model specifies applicant preferences and allows them to evolve over time. In particular, there are two periods, 0 and 1. I interpret period 0 as the currently used application period (e.g., December) and period 1 as a later point in time that can potentially be used as an alternative application period (e.g., following April).

In each period $t = 0, 1$, each applicant $a \in A$ has a strict preference relation $\succ_a^t$ over $S \cup \{\emptyset\}$, where $\emptyset$ denotes the outside option of the applicant. I distinguish $\emptyset$ and $\varnothing$, where $\emptyset$ denotes an outside option while $\varnothing$ is the empty set. $\succ_a^1$ may or may not be the same as $\succ_a^0$. $\succ_A^1$ is supposed to be a better measure of applicants' welfare than $\succ_A^0$ is since $\succ_A^1$ is at a later point in time and is likely to approximate better applicants' preferences when they finally experience educational services at assigned schools. Unless otherwise noted, I impose no restriction on the relationship between $\succ_a^0$ and $\succ_a^1$. The weak preference relation associated with $\succ_a^t$ is denoted by $\succsim_a^t$ and so I write $s \succsim_a^t \bar{s}$ if either $s \succ_a^t \bar{s}$ or $s = \bar{s}$. A preference profile of all applicants in period t is denoted by $\succ_A^t \equiv (\succ_a^t)_{a \in A}$.

Matching and Its Properties. An outcome of the model is a *matching*, which is a vector $\mu = (\mu_a)_{a \in A}$ that assigns each applicant a a seat at a school (or the outside option) $\mu_a \in A \cup \{\emptyset\}$, and where each school $s \in S$ is assigned at most q_s applicants. I denote by $\mu_s \equiv \{a \in A : s = \mu_a\}$ the set of applicants who are assigned to school s.

Let me introduce several desirable properties of a matching. I often use "w.r.t." to mean "with respect to." A matching μ is *individually rational w.r.t.* $\succ_A^t$ if $\mu_a \succsim_a^t \emptyset$ for every $a \in A$. μ is *blocked* by $(a, s) \in A \times S$ w.r.t. $(\succ_A^t, \succ_S)$ if $s \succ_a^t \mu_a$ and there exists $A' \subseteq \mu_s \cup a$ such that $A' \succ_s \mu_s$. (I denote singleton set $\{x\}$ by x when there is no room for confusion.) μ is *fair (stable) w.r.t.* $(\succ_A^t, \succ_S)$ if it is individually rational w.r.t. $\succ_A^t$ and not blocked w.r.t. $(\succ_A^t, \succ_S)$. A matching μ is *Pareto efficient for applicants w.r.t.* $\succ_A^t$ if there exists no matching μ' such that $\mu'_a \succsim_a^t \mu_a$ for all $a \in A$ and $\mu'_a \succ_a^t \mu_a$ for at least one $a \in A$. μ is *weakly Pareto efficient for applicants w.r.t.* $\succ_A^t$ if there exists no matching μ' such that $\mu'_a \succ_a^t \mu_a$ for all $a \in A$.

Mechanism and Its Properties. Given the set of applicants A and schools S, a (direct) *mechanism* is a function φ that maps each $(\succ_A^0, \succ_A^1, \succ_S)$ into a matching. Since the domain contains both $\succ_A^0$ and $\succ_A^1$, this definition allows both for static and dynamic mechanisms. The fact that NYC runs the discretionary reapplication process after the initial match suggests that it is feasible to elicit preferences at two different points in time. For example, as explained below, the current initial match mechanism in NYC is static while some of the alternatives considered in this paper are dynamic; all of them can be described as examples of this definition. A mechanism φ is *(weakly) Pareto efficient for applicants w.r.t.* $\succ_A^t$ if $\varphi(\succ_A^0, \succ_A^1, \succ_S)$ is (weakly) Pareto efficient w.r.t. $\succ_A^t$ for every $(\succ_A^0, \succ_A^1, \succ_S)$. φ is *fair (stable) w.r.t.* $(\succ_A^t, \succ_S)$ if $\varphi(\succ_A^0, \succ_A^1, \succ_S)$ is fair w.r.t. $(\succ_A^t, \succ_S)$ for every $(\succ_A^0, \succ_A^1, \succ_S)$. I introduce the following algorithm to compute a matching.

`Gale and Shapley's (Applicant-Proposing) Deferred Acceptance (DA) Algorithm.` Consider any applicant preference profile $\succ_A$, for example, $\succ_A^0$ or $\succ_A^1$. Given $(\succ_A, \succ_S)$, the (applicant-proposing) *deferred acceptance (DA) algorithm* is defined as follows.

- Step 1: Each applicant $a \in A$ applies to her most preferred acceptable school w.r.t. $\succ_a$ (if any). Each school tentatively keeps the highest-ranking applicants up to its capacity, and rejects every other applicant.

In general, for any step $t \geq 2$,

- Step t: Each applicant a who was not tentatively matched to any school in Step $(t-1)$ applies to her most preferred acceptable school w.r.t. $\succ_a$ that has not rejected her (if any). Each school tentatively keeps the highest-ranking applicants up to its capacity from the set of applicants previously tentatively matched to this school and the applicants newly applying, and rejects every other applicant.

The algorithm terminates at the first step at which no applicant applies to a school. Each applicant tentatively kept by a school at that step is allocated a seat in that school, resulting in a matching which I denote by $DA(\succ_A, \succ_S)$. It is known that $DA(\succ_A, \succ_S)$ is a fair and weakly Pareto efficient matching with respect to $(\succ_A, \succ_S)$ for any $(\succ_A, \succ_S)$ (Roth and Sotomayor, 1990; Abdulkadiroğlu and Sönmez, 2003).

As already explained in Section 2.1, the authorities in NYC ask applicants to submit their preferences around Nov and Dec. They then apply the DA algorithm to the submitted preferences to compute the matching. Recall that period 0 in my model corresponds to Nov or Dec. This leads me to interpret the current initial match mechanism as $\varphi^{DA}_{initial}(\succ_A^0, \succ_A^1, \succ_S) \equiv DA(\succ_A^0, \succ_S)$, i.e., the one that produces a matching by applying the DA algorithm to applicant preference in period 0, $\succ_A^0$. By the above properties of $DA(\succ_A, \succ_S)$ with respect to its input preferences, $\varphi^{DA}_{initial}$ has nice properties such as fairness and weak Pareto efficiency with respect to $\succ_A^0$. However, the empirical analysis in Sections 2 and 3 suggests that families' preferences for schools change from $t = 0$ to $t = 1$. Under $\succ_A^1$, therefore, $\varphi^{DA}_{initial}$ may not retain the desirable properties. In fact, I empirically show in Section 4 that $\varphi^{DA}_{initial}$ produces non-negligible efficiency losses under $\succ_A^1$.

To improve on $\varphi^{DA}_{initial}$, Section 4.2 describes two alternatives mechanisms, the *dynamic deferred acceptance mechanism* $\varphi^{DA}_{dynamic}(\succ_A^0, \succ_A^1, \succ_S) \equiv DA(\succ_A^1, \succ_S')$, where $\succ_S'$ denotes modified school priorities defined in Section 4.2. Precisely speaking, $\succ_S'$ is defined as follows. For all s,

- $a \succ_s' a'$ for any $a \in DA_s(\succ_A^0, \succ_S)$ and any $a' \notin DA_s(\succ_A^0, \succ_S)$ and

- $a'' \succ'_s a'''$ if and only if $a'' \succ_s a'''$ for any $a'', a''' \notin DA_s(\succ^0_A, \succ_S)$ or any $a'', a''' \in DA_s(\succ^0_A, \succ_S)$

The *deferred deferred acceptance mechanism* $\varphi^{DA}_{deferred}(\succ^0_A, \succ^1_A, \succ_S) \equiv DA(\succ^1_A, \succ_S)$. If $\succ^0_A = \succ^1_A$, i.e., there is no preference change, then both alternatives reduce to the current one, i.e., $\varphi^{DA}_{dynamic}(\succ^0_A, \succ^1_A, \succ_S) = \varphi^{DA}_{deferred}(\succ^0_A, \succ^1_A, \succ_S) = \varphi^{DA}_{initial}(\succ^0_A, \succ^1_A, \succ_S)$. Thus, neither of them performs worse than the current one even if there is no preference change.

The next section provides a formal characterization of the relationship among these mechanisms. In particular, I show that $\varphi^{DA}_{dynamic}$ and $\varphi^{DA}_{deferred}$ are the only "best possible" mechanisms in terms of efficiency and fairness with respect to $\succ^1_A$ and strategy-proofness. This result provides a theoretical foundation for considering these alternatives in the counterfactual analysis.

A.4.2 Result

In addition to the usual fairness and efficiency properties defined above, I introduce additional criteria to compare alternative reapplication mechanisms with $\varphi^{DA}_{initial}$, the status quo initial match mechanism currently in place in many cities including NYC.

Definition 1. A mechanism φ *always Pareto dominates* another mechanism φ' w.r.t. $\succ^t_A$ if (1) for any $(\succ^0_A, \succ^1_A, \succ_S)$ and any a, $\varphi_a(\succ^0_A, \succ^1_A, \succ_S) \succsim^t_a \varphi'_a(\succ^0_A, \succ^1_A, \succ_S)$ and (2) there are $(\succ^0_A, \succ^1_A, \succ_S)$ and a such that $\varphi_a(\succ^0_A, \succ^1_A, \succ_S) \succ^t_a \varphi'_a(\succ^0_A, \succ^1_A, \succ_S)$.

Definition 2. A mechanism φ is *less unfair* than another mechanism φ' w.r.t. $(\succ^t_A, \succ_S)$ if (1) for any $(\succ^0_A, \succ^1_A, \succ_S)$ and any (a, s), if (a, s) blocks $\varphi(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$, then (a, s) also blocks $\varphi'(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$ and (2) there are $(\succ^0_A, \succ^1_A, \succ_S)$ and (a, s) such that (a, s) blocks $\varphi'(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$ but (a, s) does *not* block $\varphi(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$.

In words, φ is less unfair than φ' if it is always the case that any applicant-school pair blocking the outcome under φ also blocks that under φ'. By definition, if φ is fair with respect to $(\succ^t_A, \succ_S)$ but φ' is not, then φ is less unfair than φ' with respect to $(\succ^t_A, \succ_S)$.

Another possible definition of being less unfair is the following: (1) for any $(\succ^0_A, \succ^1_A, \succ_S)$ and any a, if there is s such that (a, s) blocks $\varphi(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$, then there is s' such that (a, s') blocks $\varphi'(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$ (s' may or may not be the same as s) and (2) there are $(\succ^0_A, \succ^1_A, \succ_S)$ and a such that there is s' such that (a, s') blocks $\varphi'(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$ but for any s, (a, s) does *not* block $\varphi(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^t_A, \succ_S)$. The results below hold under either definition.

All properties defined so far are about fairness or efficiency with respect to stated preferences. To make sure that stated preferences reflect true ones, I would also like a mechanism to be incentive compatible.

Definition 3. A mechanism φ is *dynamically strategy-proof* if the following holds for any $(\succ_A^0, \succ_A^1, \succ_S)$, any a, and any $(\succ_a'^0, \succ_a'^1)$: $\varphi_a(\succ_A^0, \succ_A^1, \succ_S) \succsim_a^0 \varphi_a(\succ_a'^0, \succ_{-a}^0, \succ_a'^1, \succ_{-a}^1, \succ_S)$ and $\varphi_a(\succ_A^0, \succ_A^1, \succ_S) \succsim_a^1 \varphi_a(\succ_A^0, \succ_a'^1, \succ_{-a}^1, \succ_S)$.

In words, under a dynamically strategy-proof mechanism, any preference manipulation by any applicant in any period is never strictly profitable with respect to that applicant's preference in that period. I allow each applicant a to manipulate both $\succ_a^0$ and $\succ_a^1$ in $t = 0$ but only $\succ_a^1$ in $t = 1$. This restriction is justified under the interpretation that $\succ_a^0$ is already reported and fixed in $t = 1$. When φ is static, i.e., uses only $\succ_A^0$ or $\succ_A^1$ to compute the matching $\varphi(\succ_A^0, \succ_A^1, \succ_S)$, usual static strategy-proofness is the same as dynamic strategy-proofness.

In principle, I can define dynamic strategy-proofness in a more restrictive way, e.g., for any $(\succ_A^0, \succ_A^1, \succ_S)$ and any a and any $(\succ_a'^0, \succ_a'^1)$, $\varphi_a(\succ_a'^0, \succ_{-a}^0, \succ_a'^1, \succ_{-a}^1, \succ_S) \succsim_a^t \varphi_a(\succ_A^0, \succ_A^1, \succ_S)$ for both $t = 0, 1$. In the presence of potential preference reversals between $\succ_a^0$ and $\succ_a^1$, however, this definition is so restrictive that it is not satisfied even by static strategy-proof mechanisms such as $\varphi_{initial}^{DA}$ (since it is possible $\varphi_a(\succ_a'^0, \succ_{-a}^0, \succ_a'^1, \succ_{-a}^1, \succ_S) \succ_a^0 \varphi_a(\succ_A^0, \succ_A^1, \succ_S)$ but $\varphi_a(\succ_A^0, \succ_A^1, \succ_S) \succ_a^1 \varphi_a(\succ_a'^0, \succ_{-a}^0, \succ_a'^1, \succ_{-a}^1, \succ_S)$).

I regard the following five as basic desirable properties of a mechanism in my dynamic model with evolving preferences, where $\succ_a^1$ is the measure of applicants' welfare. Note that period 1 is closer to enrollment periods and the empirical analysis in the main body suggests that $\succ_a^1$ appears to be subject to less severe information frictions about schools.

(I) Fairness with respect to $(\succ_A^1, \succ_S)$
(II) Being less unfair than $\varphi_{initial}^{DA}$ with respect to $(\succ_A^1, \succ_S)$
(III) Weak Pareto efficiency with respect to $\succ_A^1$
(IV) Always Pareto dominating $\varphi_{initial}^{DA}$ with respect to $\succ_A^1$
(V) Dynamic strategy-proofness

Properties (II) and (IV), which compare an alternative mechanism with $\varphi_{initial}^{DA}$, are important since many cities including NYC are currently using $\varphi_{initial}^{DA}$ as the status quo initial match. $\varphi_{initial}^{DA}$ satisfies only (V) among (I) to (V). My goal is thus to design reapplication mechanisms that improve on $\varphi_{initial}^{DA}$ and achieve (I)-(V). The following result shows that in terms of (I)-(V), $\varphi_{dynamic}^{DA}$ and $\varphi_{deferred}^{DA}$ are the best possible mechanisms I can design.

Proposition 1. *1.A) $\varphi^{DA}_{dynamic}$ satisfies (II) being less unfair than the initial match, (III) weak Pareto efficiency, and (IV) always Pareto dominating $\varphi^{DA}_{initial}$ (call this set of desiderata A), but not others.*
1.B) $\varphi^{DA}_{deferred}$ satisfies (I) fairness, (II) being less unfair than $\varphi^{DA}_{initial}$, (III) weak Pareto efficiency, and (V) dynamic strategy-proofness (call this set of desiderata B), but not others.
2) Consider any possible mechanism φ. φ can satisfy only a subset of set A or B.

The proof is in Appendix A.4.3. Proposition 1 has several implications. First of all, recall that $\varphi^{DA}_{initial}$ does not satisfy any of desirable welfare properties (I)-(IV). Thus, $(1.A)$ and $(1.B)$ say each of $\varphi^{DA}_{dynamic}$ and $\varphi^{DA}_{deferred}$ is better than $\varphi^{DA}_{initial}$ and achieve some of welfare properties (I)-(IV) in the presence of preference changes. $(1.A)$ and $(1.B)$ also demonstrate certain tradeoffs between $\varphi^{DA}_{dynamic}$ and $\varphi^{DA}_{deferred}$. Part (2) shows that these tradeoffs are not resolvable, i.e., I cannot design a mechanism that is strictly better than any of the two mechanisms in terms of the above desiderata. In this sense, $\varphi^{DA}_{deferred}$ and $\varphi^{DA}_{dynamic}$ are *the* best possible alternatives I can obtain.

It may be useful to walk through what tradeoffs Proposition 1 embeds. A tradeoff Proposition 1 implies is that (IV) Pareto dominating $\varphi^{DA}_{initial}$ and (V) dynamic strategy-proofness are incompatible. This is a version of the classic efficiency-incentive tradeoff. There is another more important tradeoff. Proposition 1 also implies a tradeoff between (I) fairness and (IV) always Pareto dominating $\varphi^{DA}_{initial}$. This is a version of yet another classic tradeoff between efficiency and fairness.

To understand the implication of this efficiency-fairness tradeoff between (I) fairness and (IV) always Pareto dominating $\varphi^{DA}_{initial}$, let me consider a weaker version of always Pareto dominating $\varphi^{DA}_{initial}$. A mechanism φ is *never dominated* by another mechanism φ' w.r.t. $\succ^t_A$ if there is no $(\succ^0_A, \succ^1_A, \succ_S)$ such that $\varphi'(\succ^0_A, \succ^1_A, \succ_S)$ Pareto dominates $\varphi(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $\succ^t_A$. If φ dominates φ' w.r.t. $\succ^t_A$, then φ is never dominated by φ' w.r.t. $\succ^t_A$. Thus $\varphi^{DA}_{dynamic}$ is never dominated by $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$ since $\varphi^{DA}_{dynamic}$ always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$ (by Proposition 1.A). In contrast, $\varphi^{DA}_{deferred}$ is sometimes dominated by $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$ (see Lemma 3 in Appendix A.4.3). Thus, the most obvious reapplication mechanism $\varphi^{DA}_{deferred}$ (waiting until period 1 and applying the DA algorithm to applicant preferences in period 1) may make all applicants worse off *with respect to their preferences in period 1.* More generally, switching from $\varphi^{DA}_{initial}$ to $\varphi^{DA}_{deferred}$ produces "losers" who prefer $\varphi^{DA}_{initial}$ while there is no such loser in the case of $\varphi^{DA}_{dynamic}$. I empirically confirm and quantify this in Section 4.

Overall, Proposition 1 provides a theoretical basis for using $\varphi^{DA}_{deferred}$ and $\varphi^{DA}_{dynamic}$ as the best possible reapplication mechanisms in my counterfactual analysis in Section 4, which in turn empirically quantifies the effects of the two mechanisms.

A.4.3 Proof of Proposition 1

1.A) $\varphi^{DA}_{dynamic}$ always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$ since its construction always guarantees any applicant a a seat at school $DA_a(\succ^0_A, \succ_S)$ (if any) or a more preferred school w.r.t. $\succ^1_a$. It is weakly Pareto efficient w.r.t. $\succ^1_A$ since $DA(\succ_A, \succ_S)$ is weakly Pareto efficient w.r.t. input preferences $\succ_A$ for any $(\succ_A, \succ_S)$ (Abdulkadiroğlu and Sönmez, 2003). It is less unfair than $\varphi^{DA}_{initial}$ w.r.t. $(\succ^1_A, \succ_S)$ by the following Lemma 1. It violates the other two properties by Lemma 2 and Corollary 1 proven below.

Lemma 1. *$\varphi^{DA}_{dynamic}$ is less unfair than $\varphi^{DA}_{initial}$ w.r.t. $(\succ^1_A, \succ_S)$.*

Proof. Suppose (a, s) blocks $\varphi^{DA}_{dynamic}(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^1_A, \succ_S)$, i.e., $s \succ^1_a \varphi^{DA}_{dynamic,a}(\succ^0_A, \succ^1_A, \succ_S)$ and $a \succ_s a'$ for some $a' \in \varphi^{DA}_{dynamic,s}(\succ^0_A, \succ^1_A, \succ_S)$. (Note that as long as $s \succ^1_a \varphi^{DA}_{dynamic,a}(\succ^0_A, \succ^1_A, \succ_S)$, it cannot be the case $|\varphi^{DA}_{dynamic,s}(\succ^0_A, \succ^1_A, \succ_S)| < q_s$ by the construction of the DA algorithm.) It has to be the case

$$(\varphi^{DA}_{dynamic,a'}(\succ^0_A, \succ^1_A, \succ_S) =) s = DA_{a'}(\succ^0_A, \succ_S) (\equiv \varphi^{DA}_{initial,a'}(\succ^0_A, \succ^1_A, \succ_S)),$$

since otherwise $a \succ_s a'$ implies $a \succ'_s a'$ and thus (a, s) blocks $\varphi^{DA}_{dynamic}(\succ^0_A, \succ^1_A, \succ_S) \equiv DA(\succ^1_A, \succ'_S)$ w.r.t. $(\succ^1_A, \succ'_S)$, a contradiction to the fact that no (a, s) blocks $DA(\succ_A, \succ_S)$ w.r.t. $(\succ_A, \succ_S)$ for any $(\succ_A, \succ_S)$. Then

$$s \succ^1_a \varphi^{DA}_{dynamic,a}(\succ^0_A, \succ^1_A, \succ_S) \succeq^1_a \varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S),$$

where the second weak preference is because $\varphi^{DA}_{dynamic}$ always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$. Combined with $\varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S) \equiv DA_{a'}(\succ^0_A, \succ_S) = s$, this means (a, s) also blocks $\varphi^{DA}_{initial}(\succ^0_A, \succ^1_A, \succ_S)$ w.r.t. $(\succ^1_A, \succ_S)$. □

1.B) $\varphi^{DA}_{deferred} \equiv DA(\succ^1_A, \succ_S)$ is fair w.r.t. $(\succ^1_A, \succ_S)$ and weakly Pareto efficient w.r.t. $\succ^1_A$ since $DA(\succ_A, \succ_S)$ has these properties w.r.t. any input preferences (Abdulkadiroğlu and Sönmez, 2003). Since $\varphi^{DA}_{initial}$ is not fair w.r.t. $(\succ^1_A, \succ_S)$, $\varphi^{DA}_{deferred}$'s fairness implies that $\varphi^{DA}_{deferred}$ is less unfair than $\varphi^{DA}_{initial}$ w.r.t. $(\succ^1_A, \succ_S)$. $\varphi^{DA}_{deferred}$ is dynamically strategy-proof since $\varphi^{DA}_{deferred}$ is static and statically strategy-proof while for static mechanisms, dynamic strategy-proofness is equivalent to static strategy-proofness. $\varphi^{DA}_{deferred}$ does not always Pareto dominate $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$ by Corollary 1 shown below.

2) This is implied by Lemma 2 and Corollary 1.

Lemma 2. *There is no mechanism that always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$ and is dynamically strategy-proof.*

Proof. Suppose to the contrary that there is a mechanism φ that satisfies the above two properties. Since φ always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$, there exist $(\succ^0_A, \succ^1_A, \succ_S)$ and a such that $\varphi_a(\succ^0_A, \succ^1_A, \succ_S) \neq \varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S)$.

Claim 1. There exist $(\succ^0_A, \succ^1_A, \succ_S)$ and a such that $\varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S) \succ^0_a \varphi_a(\succ^0_A, \succ^1_A, \succ_S$).

Proof. Otherwise, for any $(\succ^0_A, \succ^1_A, \succ_S)$ and any a with $\varphi_a(\succ^0_A, \succ^1_A, \succ_S) \neq \varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S)$, it is the case that $\varphi_a(\succ^0_A, \succ^1_A, \succ_S) \succ^0_a \varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S)$. This means that φ always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^0_A$. Consider $\succ'^0_a$: $\varphi_a(\succ^0_A, \succ^1_A, \succ_S), \emptyset$. Since $\varphi^{DA}_{initial}$ is strategy-proof, $\varphi^{DA}_{initial,a}(\succ'^0_a, \succ^0_{-a}, \succ^1_A, \succ_S) = \emptyset$. Since φ always Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^0_A$, the CLAIM in Abdulkadiroğlu et al. (2009) implies that $\varphi_a(\succ'^0_a, \succ^0_{-a}, \succ^1_A, \succ_S) = \emptyset$. This means that $\succ^0_a$ is a profitable deviation for a w.r.t. $\succ'^0_a$ under φ when the true preference profile is $(\succ'^0_a, \succ^0_{-a}, \succ^1_A, \succ_S)$, a contradiction to the assumption that φ is dynamically strategy-proof. Thus, there exist $(\succ^0_A, \succ^1_A, \succ_S)$ and a such that $\varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S) \succ^0_a \varphi_a(\succ^0_A, \succ^1_A, \succ_S)$. □

For any a such that $\varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S) \succ^0_a \varphi_a(\succ^0_A, \succ^1_A, \succ_S)$, who exists by Claim 1, consider the following preference: $\succ'^1_a$: $\varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S), \emptyset$. Since $\varphi_a(\succ^0_A, \succ'^1_a, \succ^1_{-a}, \succ_S) = \varphi^{DA}_{initial,a}(\succ^0_A, \succ^1_A, \succ_S) \succ^0_a \varphi_a(\succ^0_A, \succ^1_A, \succ_S)$, where the first equality is by the assumption that φ always dominates $\varphi^{DA}_{initial}$. $\succ'^1_a$ is a profitable deviation for a w.r.t. $\succ^0_a$ at $(\succ^0_A, \succ^1_A, \succ_S)$, a contradiction. □

Definition 4. A school preference profile $\succ_S$ is *acyclic* if there exist no $s_1, s_2 \in S$ and $a_1, a_2, a_3 \in A$ such that

- $a_1 \succ_{s_1} a_2 \succ_{s_1} a_3 \succ_{s_2} a_1$ and
- there exist (possibly empty) disjoint sets of students $A_{s_1}, A_{s_2} \subseteq A \setminus \{a_1, a_2, a_3\}$ such that $|A_{s_1}| = q_{s_1} - 1, |A_{s_2}| = q_{s_2} - 1$, $a \succ_{s_1} a_2$ for every $a \in A_{s_1}$ and $a' \succ_{s_2} a_1$ for every $a' \in A_{s_2}$.

Lemma 3. *There is a mechanism that is fair w.r.t. $(\succ^1_A, \succ_S)$ and never dominated by $\varphi^{DA}_{initial}$ if and only if $\succ_S$ is acyclic.*

Proof. Acyclicity is sufficient because acyclicity guarantees that $\varphi^{DA}_{deferred}$ is Pareto efficient for applicants w.r.t. $\succ^1_A$ (Ergin, 2002) and so is not dominated by $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$. For

the necessity part, suppose to the contrary that though $\succ_S$ is cyclic, a mechanism that is fair w.r.t. $(\succ_A^1, \succ_S)$ is never dominated by $\varphi_{initial}^{DA}$. By the definition of acyclicity, there exist $s_1, s_2 \in S$ and $a_1, a_2, a_3 \in A$ such that $a_1 \succ_{s_1} a_2 \succ_{s_1} a_3 \succ_{s_2} a_1$ and there exist (possibly empty) disjoint sets of applicants $A_{s_1}, A_{s_2} \subseteq A \setminus \{a_1, a_2, a_3\}$ such that $|A_{s_1}| = q_{s_1} - 1, |A_{s_2}| = q_{s_2} - 1$, $a \succ_{s_1} a_2$ for every $a \in A_{s_1}$ and $a' \succ_{s_2} a_1$ for every $a' \in A_{s_2}$. Consider the following preference profile $\succ_A^t$ of applicants (With 3 or more schools, it is easy to expand this example so that a_2 ranks some school in period 0):

$$
\begin{aligned}
&\succ_{a_1}^t (t = 0, 1) : s_2, s_1, \emptyset, \\
&\succ_{a_2}^0 : \emptyset, \\
&\succ_{a_2}^1 : s_1, \emptyset, \\
&\succ_{a_3}^t (t = 0, 1) : s_1, s_2, \emptyset, \\
&\succ_l^t (t = 0, 1) : s_1, \emptyset, \forall l \in A_{s_1}, \\
&\succ_m^t (t = 0, 1) : s_2, \emptyset, \forall m \in A_{s_2}, \\
&\succ_n^t (t = 0, 1) : \emptyset, \forall n \in A \setminus (A_{s_1} \cup A_{s_2} \cup \{a_1, a_2, a_3\}).
\end{aligned}
$$

$\varphi_{initial}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ matches $\{a_3\} \cup A_{s_1}$ to s_1, $\{a_1\} \cup A_{s_2}$ to s_2, and leaves all the other applicants unmatched. $\varphi_{deferred}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ matches $\{a_1\} \cup A_{s_1}$ to s_1, $\{a_3\} \cup A_{s_2}$ to s_2, and leaves all the other students unmatched. However, since $s_2 \succ_{a_1}^t s_1$ and $s_1 \succ_{a_3}^t s_2$ for $t = 0, 1$, $\varphi_{deferred}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ is Pareto dominated by $\varphi_{initial}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$ by the applicants w.r.t. both $\succ_A^0$ and $\succ_A^1$. Since any mechanism that is fair w.r.t. $(\succ_A^1, \succ_S)$ is (weakly) Pareto dominated by $\varphi_{deferred}^{DA}(\succ_A^0, \succ_A^1, \succ_S)$, this implies that any mechanism that is fair w.r.t. $(\succ_A^1, \succ_S)$ is Pareto dominated by $\varphi_{initial}^{DA}$ w.r.t. both $\succ_A^0$ and $\succ_A^1$, a contradiction. □

Corollary 1. *There is no mechanism that is fair w.r.t.* $(\succ_A^1, \succ_S)$ *and Pareto dominates* $\varphi_{initial}^{DA}$ *w.r.t.* $\succ_A^1$.

A.4.4 Extension

Proposition 1 is based on Definition 3 of dynamic strategy-proofness. Definition 3 may appear to be restrictive in that it allows applicants to jointly manipulate and commit to not only $\succ_a^0$ but also $\succ_a^1$ in period 0. The following alternative, weaker definition of dynamic strategy-proofness excludes such joint manipulations.

Definition 5. A mechanism φ is *weakly dynamically strategy-proof* if the following holds for any $(\succ_A^0, \succ_A^1, \succ_S)$ and any a: (0) for any $\succ_a'^0$, $\varphi_a(\succ_a'^0, \succ_{-a}^0, \succ_A^1, \succ_S) \succsim_a^0 \varphi_a(\succ_A^0, \succ_A^1, \succ_S)$ and (1) for any $\succ_a'^1$, $\varphi_a(\succ_A^0, \succ_a'^1, \succ_{-a}^1, \succ_S) \succsim_a^1 \varphi_a(\succ_A^0, \succ_A^1, \succ_S)$.

In words, under a weakly dynamically strategy-proof mechanism, any myopic or one-shot preference manipulation by any applicant in any period is never strictly profitable with respect to that applicant's preference in that period. Weak dynamic strategy-proofness thus requires a mechanism to be always immune to any myopic manipulation. Weak dynamic strategy-proofness is implied by dynamic strategy-proofness because the former allows for a smaller set of potential manipulations than the latter. Parts (1.A) and (1.B) of Proposition 1 remain to hold even under weak dynamic strategy-proofness (instead of dynamic strategy-proofness). It is open whether the remaining part (2) of Proposition 1 also remains correct. Nevertheless, the following partial result is true.

Lemma 4. *There is no mechanism that always dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$, is weakly dynamically strategy-proof, and is weakly Pareto efficient w.r.t. $\succ^1_A$.*

The proof is in Appendix A.4.5. Lemma 4 and the proof of Proposition 1 imply that the only potentially necessary modification to Proposition 1 when using weak dynamic strategy-proofness is that there may exist a mechanism that always dominates $\varphi^{DA}_{initial}$ with respect to $\succ^1_A$, is weakly dynamically strategy-proof, but is *not* weakly Pareto efficient w.r.t. $\succ^1_A$ (and not fair with respect to $\succ^1_A$). I do not consider such a mechanism in the counterfactual analysis in Section 4 since even if it exists and I can construct it, the counterfactual analysis is mainly interested in efficiency gains over $\varphi^{DA}_{initial}$ and so prefers $\varphi^{DA}_{dynamic}$ (which is weakly Pareto efficient w.r.t. $\succ^1_A$) over it (which has to be not weakly Pareto efficient w.r.t. $\succ^1_A$).

A.4.5 Proof of Lemma 4

Suppose to the contrary that there is a mechanism φ that satisfies the above three properties. Consider a problem with four applicants a_1, a_2, a_3, a_4 and four schools s_1, s_2, s_3, s_4 each of which has the capacity of one. Their preferences are as follows:

$\succ_{s_1}$: Anything with $a_3 \succ_{s_1} a_4 \succ_{s_1} a_1$
$\succ_{s_2}$: Anything
$\succ_{s_3}$: Anything with $a_1 \succ_{s_3} a_3$
$\succ_{s_4}$: Anything with $a_4 \succ_{s_4} a_2$
$\succ^0_{a_1}$: $s_1, s_2, s_3, ...$
$\succ^0_{a_2}$: $s_4, s_2, ...$
$\succ^0_{a_3}$: $s_3, s_1, ...$
$\succ^0_{a_4}$: $s_1, s_4, ...$
$\succ^1_{a_1}$: $s_1, s_2, s_3, ...$
$\succ^1_{a_2}$: $s_3, s_4, s_2, ...$

$\succ^1_{a_3}$: $s_3, s_2, s_1, ...$
$\succ^1_{a_4}$: $s_3, s_4, s_1, ...$

Note that $\varphi^{DA}_{initial}(\succ^0_A, \succ^1_A, \succ_S) = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ s_2 & s_4 & s_3 & s_1 \end{pmatrix}$ and this matching is Pareto efficient w.r.t. $\succ^1_A$. By the assumption that φ always dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$, $\varphi(\succ^0_A, \succ^1_A, \succ_S) = \varphi^{DA}_{initial}(\succ^0_A, \succ^1_A, \succ_S) = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ s_2 & s_4 & s_3 & s_1 \end{pmatrix}$. Thus $\varphi_{a_1}(\succ^0_A, \succ^1_A, \succ_S) = s_2$.

Now let me consider the following preference of a_1:

$\succ'^0_{a_1}$: $s_3, ...$

Note that $\varphi^{DA}_{initial}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ s_3 & s_2 & s_1 & s_4 \end{pmatrix}$. This matching is not weakly Pareto efficient w.r.t. $\succ^1_A$ since it is strongly Pareto dominated by $\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ s_1 & s_4 & s_2 & s_3 \end{pmatrix}$. By the assumption that φ is weakly Pareto efficient w.r.t. $\succ^1_A$ and always dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^1_A$, $\varphi(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$ has to be a matching that is weakly Pareto efficient w.r.t. $\succ^1_A$ (and so $\varphi(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) \neq \varphi^{DA}_{initial}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$) and Pareto dominates $\varphi^{DA}_{initial}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$. I use the following fact.

Lemma 5. $\varphi_{a_1}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) = s_1$.

Proof. By construction of $\succ^1_{a_1}$, $\varphi_{a_1}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ_S) = s_1$ or s_2 or s_3. It is thus enough to show $\varphi_{a_1}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) \neq s_2$ or s_3.

First, $\varphi_{a_1}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) \neq s_3$ by the following reason: Suppose to the contrary that $\varphi_{a_1}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) = s_3$. Inspections show that there is no matching μ such that $\mu_{a_1} = s_3$ and μ Pareto dominates $\varphi^{DA}_{initial}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$ with respect to $\succ^1_A$ [37]; thus $\varphi_{a_1}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) = s_3$ implies that $\varphi(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$ does not Pareto dominates

[37] Suppose to the contrary that there is some matching μ such that $\mu_{a_1} = s_3$ and μ Pareto dominates $\varphi^{DA}_{initial}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$ with respect to $\succ^1_A$.

<u>*Case 1*</u>: If $\mu_{a_4} \succ^1_{a_4} \varphi^{DA}_{initial,a_4}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) = s_4$, then $\mu_{a_4} = s_3$, which is the only school a_4 strictly prefers to $s_4 = \varphi^{DA}_{initial,a_4}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$. But this contradicts the assumption of $\mu_{a_1} = s_3$.

<u>*Case 2*</u>: If $\mu_{a_3} \succ^1_{a_3} \varphi^{DA}_{initial,a_3}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S) = s_1$, then $\mu_{a_3} = s_2$, since s_2 and s_3 are the only schools a_3 strictly prefers to $s_1 = \varphi^{DA}_{initial,a_3}(\succ'^0_{a_1}, \succ^0_{-a_1}, \succ^1_A, \succ_S)$ w.r.t. $\succ^1_{a_3}$ and $\mu_{a_1} = s_3$. Since $s_2 = \varphi^{DA}_{initial,a_2}(\succ^0_A, \succ'^1_{a_1}, \succ'^1_{a_1}, \succ^1_{-a_1}, \succ_S)$, it has to be the case $\mu_{a_2} = s_4$: s_3 and s_4 are the only schools a_2 strictly prefers to $s_2 = \varphi^{DA}_{initial,a_2}(\succ^0_A, \succ'^1_{a_1}, \succ^1_A, \succ_S)$ and $\mu_{a_1} = s_3$ by assumption. Then, since $s_4 = \varphi^{DA}_{initial,a_4}(\succ^0_A, \succ'^1_{a_1}, \succ^1_A, \succ_S)$, it has to be the case that $\mu_{a_4} \succ^1_{a_4} \varphi^{DA}_{initial,a_4}(\succ^0_A, \succ'^1_{a_1}, \succ^1_A, \succ_S)$. Thus this case reduces to Case 1, a contradiction.

$\varphi^{DA}_{initial}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S)$, which contradicts the assumption that φ is weakly Pareto efficient and always dominates $\varphi^{DA}_{initial}$ w.r.t. $\succ^{1}_{A}$.

Next, $\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) \neq s_2$ by the following reason: Suppose to the contrary that $\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) = s_2$. Consider the following preference of a_1:

$\succ'^{1}_{a_1}$: $s_1, s_3, ...$

Note that $\varphi^{DA}_{initial}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ s_3 & s_2 & s_1 & s_4 \end{pmatrix}$. This matching is not weakly Pareto efficient w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$ since it is strongly Pareto dominated by $\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ s_1 & s_4 & s_2 & s_3 \end{pmatrix}$ w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$. By the assumption that φ is weakly Pareto efficient w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$ and always dominates $\varphi^{DA}_{initial}$ w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$, $\varphi(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ has to be a matching that is weakly Pareto efficient w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$ (and so $\varphi(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) \neq \varphi^{DA}_{initial}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$) and Pareto dominates $\varphi^{DA}_{initial}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$. This implies $\varphi_{a_1}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_1$ by the following reason: Inspections show that there is no matching μ such that $\mu_{a_1} = s_3$ and μ Pareto dominates $\varphi^{DA}_{initial}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$ [38]; thus $\varphi_{a_1}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_3$ (the only other match for a_1 than s_1 by the construction of $\succ'^{1}_{a_1}$ and $\varphi^{DA}_{initial,a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$) implies that $\varphi(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ does not Pareto dominates $\varphi^{DA}_{initial}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S$

<u>*Case 3*</u>: If $\mu_{a_2} \succ^{1}_{a_2} \varphi^{DA}_{initial,a_2}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) = s_2$, then $\mu_{a_2} = s_4$, since s_3 and s_4 are the only schools a_2 strictly prefers to $s_2 = \varphi^{DA}_{initial,a_2}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S)$ w.r.t. $\succ^{1}_{a_2}$ and $\mu_{a_1} = s_3$. Then, since $s_4 = \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ^{1}_{A}, \succ_S)$, it has to be the case that $\mu_{a_4} \succ^{1}_{a_4} \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ^{1}_{A}, \succ_S)$. Thus this case reduces to Case 1, a contradiction.

[38] Suppose to the contrary that there is some matching μ such that $\mu_{a_1} = s_3$ and μ Pareto dominates $\varphi^{DA}_{initial}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$.

<u>*Case 1*</u>: If $\mu_{a_4} \succ^{1}_{a_4} \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_4$, then $\mu_{a_4} = s_3$, the only school a_4 strictly prefers to $s_4 = \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ w.r.t. $\succ^{1}_{a_4}$. But this contradicts the assumption of $\mu_{a_1} = s_3$.

<u>*Case 2*</u>: If $\mu_{a_3} \succ^{1}_{a_3} \varphi^{DA}_{initial,a_3}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_1$, then $\mu_{a_4} = s_2$, since s_2 and s_3 are the only schools a_3 strictly prefers to $s_1 = \varphi^{DA}_{initial,a_3}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ w.r.t. $\succ^{1}_{a_3}$ and $\mu_{a_1} = s_3$ by assumption. Since $s_2 = \varphi^{DA}_{initial,a_2}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$, it has to be the case $\mu_{a_2} = s_4$: s_3 and s_4 are the only schools a_2 strictly prefers to $s_2 = \varphi^{DA}_{initial,a_2}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ and $\mu_{a_1} = s_3$ by assumption. Then, since $s_4 = \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$, it has to be the case that $\mu_{a_4} \succ^{1}_{a_4} \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$. Thus this case reduces to Case 1, a contradiction.

<u>*Case 3*</u>: If $\mu_{a_2} \succ^{1}_{a_2} \varphi^{DA}_{initial,a_2}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_2$, then $\mu_{a_4} = s_4$, since s_3 and s_4 are the only schools a_2 strictly prefers to $s_2 = \varphi^{DA}_{initial,a_2}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$ w.r.t. $\succ^{1}_{a_2}$ and $\mu_{a_1} = s_3$ by assumption. Since $s_4 = \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$, it has to be the case $\mu_{a_4} \succ^{1}_{a_4} \varphi^{DA}_{initial,a_4}(\succ^{0}_{A}, \succ'^{1}_{a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S)$. Thus this case reduces to Case 1, a contradiction.

), which contradicts the assumption that φ Pareto dominates $\varphi^{DA}_{initial}$ w.r.t. $(\succ'^{1}_{a_1}, \succ^{1}_{-a_1})$. Thus $\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_1$, but this means

$$\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ'^{1}_{a_1}, \succ^{1}_{-a_1}, \succ_S) = s_1 \succ^{1}_{a_1} s_2 = \varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S),$$

a contradiction to the assumption that φ is dynamically strategy-proof. Therefore it has to be the case that $\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) \neq s_2$.

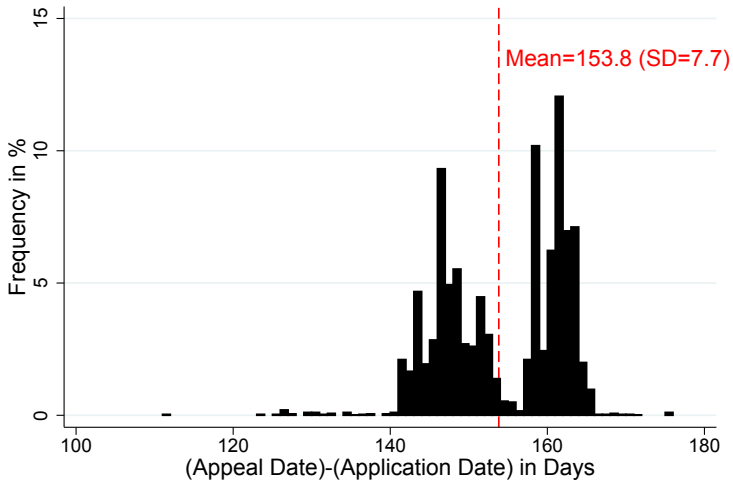
Recall that by construction of $\succ^{1}_{a_1}$, $\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) = s_1$ or s_2 or s_3. Thus the above discussions imply $\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) = s_1$. □

Lemma 4 and the paragraph right after the preference descriptions show

$$\varphi_{a_1}(\succ'^{0}_{a_1}, \succ^{0}_{-a_1}, \succ^{1}_{A}, \succ_S) = s_1 \succ^{t}_{a_1} s_2 = \varphi_{a_1}(\succ^{0}_{A}, \succ^{1}_{A}, \succ_S) \text{ for } t = 0, 1,$$

which contradicts the assumption that φ is dynamically strategy-proof. Note that this proof shows a stronger result than necessary in that $\succ'^{0}_{a_1}$ is a profitable manipulation w.r.t. both $\succ^{0}_{a_1}$ and $\succ^{1}_{a_1}$.

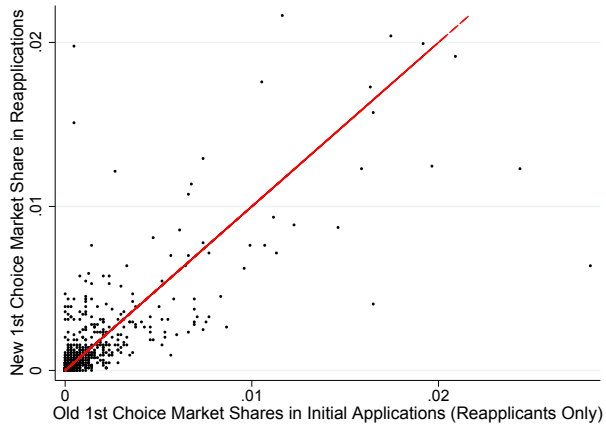
Figure A.1: Timeline of the First-round and Reapplication Process: Details



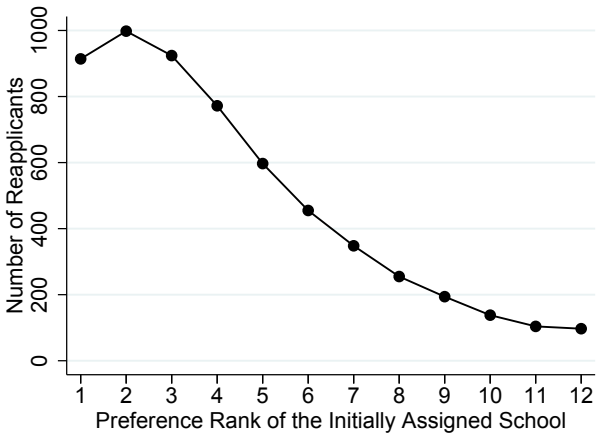
Notes: This figure draws the histogram of the difference between the date at which the initial application is filed and the date at which the reapplication is filed (both are conditional on reapplicants). See Section 2.1 for discussions about this figure.

Figure A.2: Evolving School Choices: Details

(a) Changing Market Shares of Schools



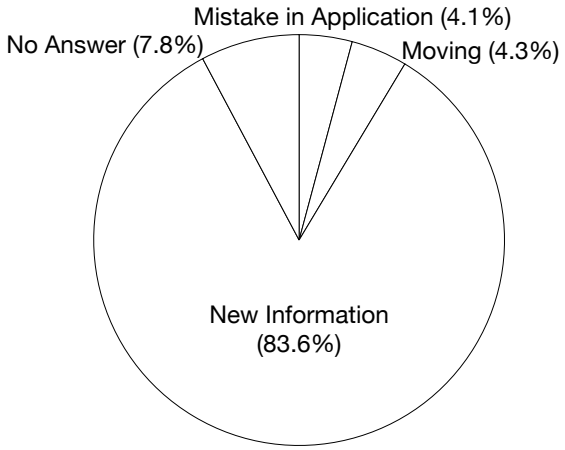
(b) Reapplications After Getting Top Choices



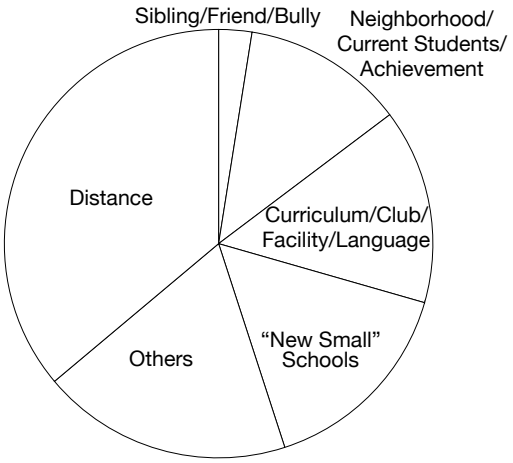
Notes: Panel A.2a plots the first choice market share of each school in initial applications (the x axis) and reapplicants (the y axis), where the *first choice market share* of a school is defined as the fraction of applicants who rank it first among all applicants who make a first choice. I compute both the old and new shares conditional on reapplicants. Panel A.2b correlates the number of reapplicants with the preference rank of the initially assigned school with respect to the initial preference. See Section 2.1 for discussions about this figure.

Figure A.3: Self-reported Reasons for Reapplication: Details

(a) Self-reported Reasons for Reapplication: Reapplicants with Choice Reversals



(b) Breakdown of "New Information"



Notes: Panel A.3a classifies self-reported reasons for reapplication conditional on applicants who reapply and exhibit choice reversals. Panel A.3b focuses on the "new information" category in Panel A.3a and breaks it down into sub-categories. "Distance" is different from "Moving After Application" in that the former does not refer to any address change. See also Section 2.2 for discussions about this figure.

Table A.1: Learning about Schools: Details

	R2 change from initial applications to reapplications for reapplicants						
	Grade 7 reading grade category			Race			
	High	Middle	Low	White	Asian	Hispanic	Black
Specification 1 (Location/borough dummies)	+45%	+32%	-30%	+2%	+54%	+46%	+8%
Specification 2 (1+Academic performance dummy)	+61%	+37%	+10%	+11%	+71%	+6%	+25%
Specification 3 (2+Program type dummies)	+102%	+105%	+201%	+64%	+125%	+112%	+68%
Specification 4 (3+Capacity dummies)	+41%	+32%	+58%	+35%	+52%	+28%	+32%

Notes: This table shows the last column in Table 4 conditional on each demographic group. The details of included characteristics are in Appendix A.2. See also Section 2.2 for discussions about this table.

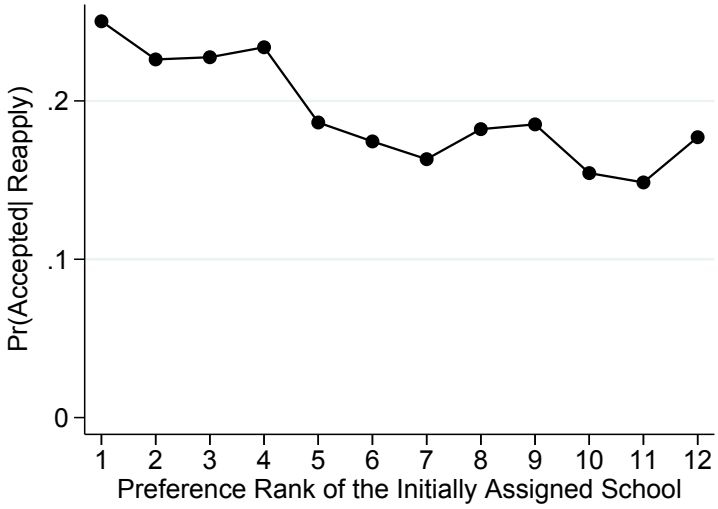
Table A.2: Growing Response of Choices to Distance and Academic Achievement: Details

		Grade 7 reading grade category								
		High			Middle			Low		
		Initial applications	Reapplications	Change	Initial applications	Reapplications	Change	Initial applications	Reapplications	Change
			Reapplicants			Reapplicants			Reapplicants	
Distance (in miles)	1st choice school	5.0	4.0	-20.3%	5.0	3.9	-20.7%	4.7	3.6	-23.8%
	Schools ranked above initial assignment	5.2	4.1	-21.2%	5.2	4.0	-21.9%	4.8	3.7	-23.1%
	All ranked schools	5.3	4.1	-22.6%	5.3	4.0	-23.9%	5.0	3.7	-25.8%
Schools with low academic performance	1st choice school	3.4%	3.5%	4.7%	11.7%	10.7%	-9.0%	17.1%	17.3%	1.3%
	Schools ranked above initial assignment	2.4%	3.6%	52.6%	10.3%	10.7%	4.0%	17.3%	18.4%	6.0%
	All ranked schools	6.6%	3.6%	-44.7%	14.6%	10.7%	-26.4%	19.3%	18.4%	-4.8%

		Race					
		White			Asian		
		Initial applications	Reapplications	Change	Initial applications	Reapplications	Change
			Reapplicants			Reapplicants	
Distance (in miles)	1st choice school	4.3	3.6	-17.7%	4.6	3.4	-25.3%
	Schools ranked above initial assignment	4.3	3.6	-16.4%	4.8	3.5	-26.4%
	All ranked schools	4.4	3.6	-17.8%	4.9	3.5	-27.9%
Schools with low academic performance	1st choice school	4.2%	4.0%	-5.3%	4.3%	4.2%	-2.0%
	Schools ranked above initial assignment	3.4%	4.0%	19.2%	3.9%	4.3%	8.9%
	All ranked schools	6.3%	4.0%	-36.1%	6.9%	4.3%	-38.4%
		Hispanic			Black		
		Initial applications	Reapplications	Change	Initial applications	Reapplications	Change
			Reapplicants			Reapplicants	
Distance (in miles)	1st choice school	4.8	3.8	-21.8%	5.5	4.4	-19.6%
	Schools ranked above initial assignment	5.0	3.9	-22.9%	5.7	4.5	-21.2%
	All ranked schools	5.1	3.9	-24.4%	5.9	4.5	-24.0%
Schools with low academic performance	1st choice school	16.8%	15.0%	-10.6%	11.0%	10.9%	-0.5%
	Schools ranked above initial assignment	14.9%	15.3%	2.5%	10.4%	10.7%	3.1%
	All ranked schools	19.0%	15.3%	-19.4%	15.2%	10.7%	-29.5%

Notes: This table shows the average distance to and academic achievement level of ranked schools in initial applications and reapplications, conditional on each demographic group. *Schools with low academic performance* are "schools in need of improvement" in the official school brochure issued by NYC. Under the No Child Left Behind Act, New York State establishes annual performance goals in Mathematics and English Language Arts for all NYC public schools. Schools that do not meet these goals for two consecutive years are identified as schools in need of improvement. See Section 2.2 for discussions about this table.

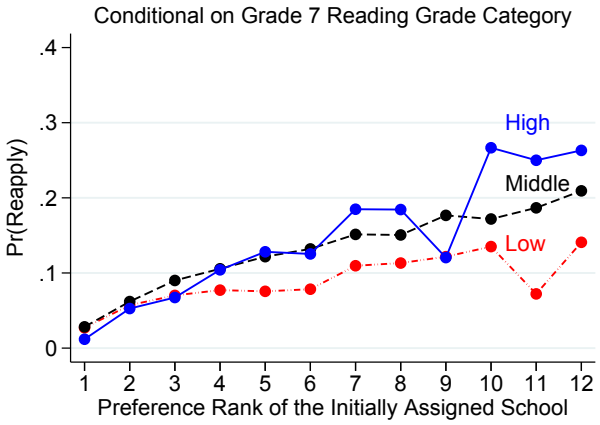
Figure A.4: Empirical Probability of Reapplication Acceptance



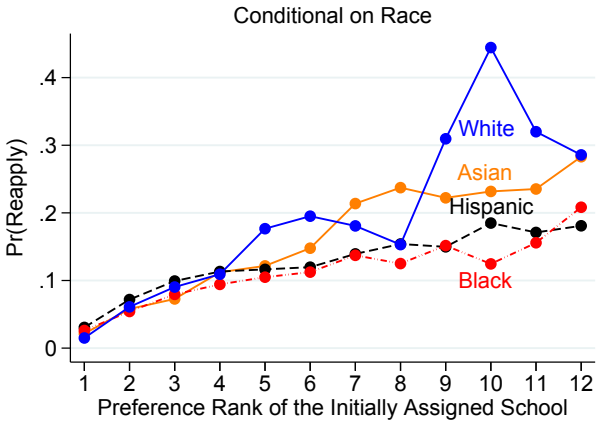
Notes: Panel A.4 correlates the probability of reapplication acceptance (being accepted by some school) conditional on reapplying with the preference rank of the initially assigned school with respect to the initial preference. See Section 3.2 for discussions about this figure.

Figure A.5: Robustness & Heterogeneity of Demand Changes & Switching Costs

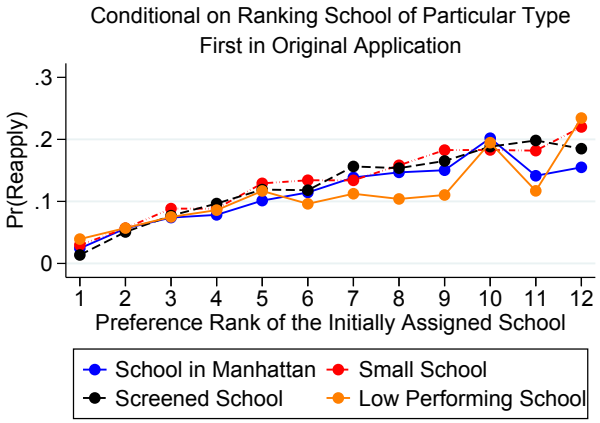
(a) Heterogeneity across Baseline Test Scores



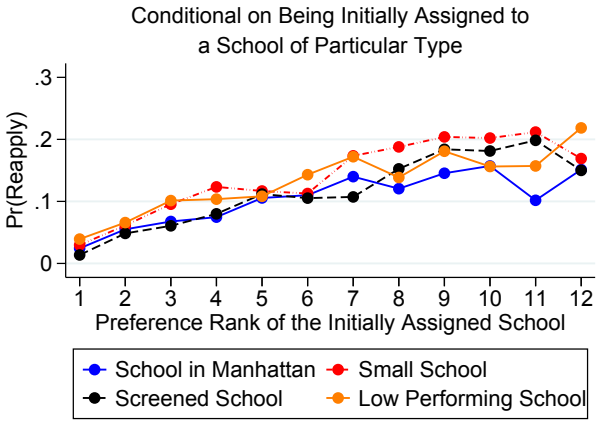
(b) Heterogeneity across Races



(c) *Homo*geneity across Old 1st Choice Schools



(d) *Homo*geneity across Initial Assignments



Notes: These figures correlate the conditional probability of reapplying to the preference rank of the initially assigned school with respect to the initial preference conditional on a group defined by a baseline characteristic. Combined with the logic in Figure 4 and Section 3.2, they suggest the (absence of) heterogeneity of demand changes and switching costs across groups. The construction of the groups are in Appendix A.1. See Section 3.2 for discussions about these figures.

Table A.3: Covariate Balance Between Applicants with and without First Choice Offers

	Covariate balance coefficients	
	Causal OLS	Descriptive OLS
Female	0.015	-0.003
	(0.011)	(0.003)
Living area		
Manhattan	-0.002	-0.007***
	(0.005)	(0.002)
Brooklyn	-0.005	0.005**
	(0.005)	(0.002)
Queens	0.003	-0.047***
	(0.004)	(0.003)
Bronx	0.004	0.007**
	(0.004)	(0.003)
Staten Island	0.000	0.041***
	(0.001)	(0.002)
Home language		
Spanish	-0.006	0.001
	(0.005)	(0.003)
Chinese	0.002	0.004***
	(0.010)	(0.001)
N (students)	11265	90978
Ethnicity		
Asian	-0.008*	-0.002
	(0.005)	(0.002)
Black	0.007	-0.032***
	(0.010)	(0.003)
Hispanic	0.000	0.005
	(0.010)	(0.003)
White	-0.008	0.052***
	(0.006)	(0.002)
1st choice school & marginal priority controls	Yes	No
N (students)	10,369	82,301
N (schools)	224	677

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%

Notes: This table reports estimates of the causal effect of being assigned to the first choice school on baseline covariates based on the linear probability model described in the main text. "Descriptive OLS" is the regression of the reapplication dummy on the first choice assignment dummy with no control or sample restriction. Standard errors are in parentheses. See Section 3.2 and Appendix A.3.1 for discussions about this table.

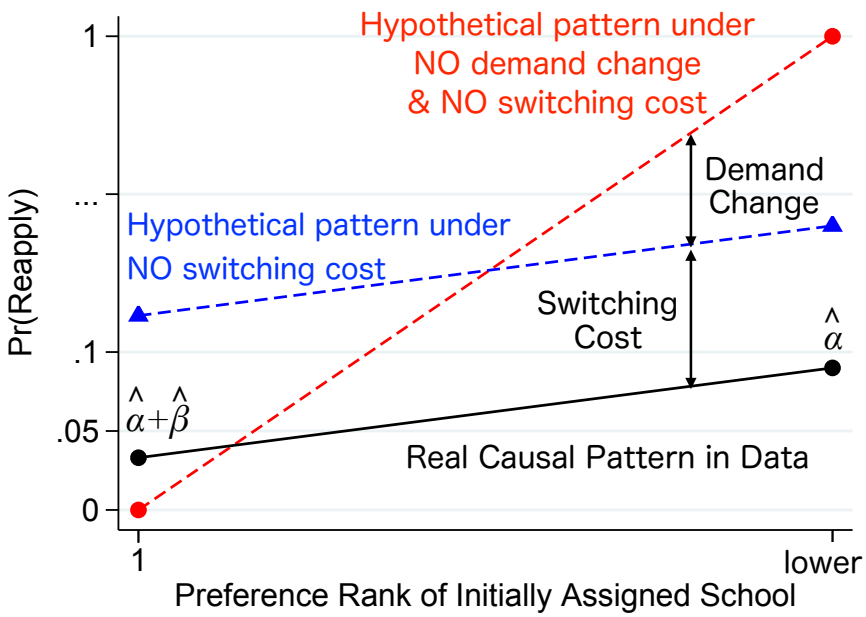
Table A.4: Causal Effect of Being Assigned to the First Choice School on Reapplying

	Dependent variable: Reapplication dummy			
	Causal OLS	Causal probit marginal effect	Causal logit marginal effect	Descriptive OLS
Coefficient on offer from 1st choice school	-0.056***	-0.060***	-0.057***	-0.070***
	(0.005)	(0.005)	(0.005)	(0.002)
1st choice school & marginal priority controls		Yes		No
N (students)		11265		90978
N (schools)		224		677

***: Significant at 1%

Notes: This table reports estimates of the causal effect of being assigned to the first choice school on the probability of reapplying based on the linear probability, probit, and logit models described in the main text. "Descriptive OLS" is the regression of the reapplication dummy on the first choice assignment dummy with no control or sample restriction. Standard errors are in parentheses. See Section 3.2 and Appendix A.3.1 for discussions about this table.

Figure A.6: Evidence of Demand Changes and Switching Costs: Structural Version



Notes: This figure illustrates how to separately identify switching costs and demand changes from the solid black line observed in the data. The solid black line correlates the conditional probability of reapplying to the preference rank of the initially assigned school with respect to the initial preference conditional on applicants who are randomly assigned to their first and lower choice schools. As detailed in the main text, $\hat{\alpha}+\hat{\beta}$ and $\hat{\alpha}$ are estimates of the conditional probabilities of reapplying conditional on being assigned to the first choice school and a non-first-choice school, respectively, in $\cup_s First_s$, where applicants are randomly assigned between the first choice school and a non-first-choice school. Combined with the logic in Figure 4 and Section 3.2, this suggests the presence of both switching costs and demand changes. See Section 3.2 and Appendix A.3.1 for discussions about this figure.

Table A.5: Tests of Switching Costs and Demand Changes

Behavioral hypothesis	Statistical hypothesis	Wald test result	p value
H_0: No switching costs & no demand changes	H_0: $\alpha+\beta=0$	Rejected	<0.01%
	H_0: $\alpha=1$	Rejected	<0.01%
H_0: No switching costs	H_0: $\alpha+\beta \geq 1-\alpha$	Rejected	<0.01%

Notes: This table reports results of Wald tests of the null hypothesis that there are no switching costs or the hypothesis that there are neither switching costs nor demand changes. See Appendix A.3.1 for discussions about this table.

Table A.6: Estimates from the Rational Expectation Model: Preferences and Frictions

	Black or hispanic								White or asian							
	Grade 7 reading grade category: high				Low or middle				High				Low or middle			
	$\beta a\,(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$		$\beta a\,(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$		$\beta a\,(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$		$\beta a\,(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Negative distance (-10 miles, log-normal for $\beta a\,(1+fa)$)	0.90***	0.60***	4.52***	1.77	1.08***	0.59***	13.14***	6.14***	1.85***	0.74***	5.02***	1.94***	1.46***	0.66***	5.66***	0.70***
	(0.02)	(0.01)	(1.39)	(1.34)	(0.01)	(0.02)	(3.60)	(1.76)	(0.20)	(0.05)	(0.91)	(0.44)	(0.01)	(0.01)	(0.19)	(0.19)
School type unscreened dummy	-0.56*	1.27***	-2.52*	2.11**	-0.75	1.02	-2.88	5.00***	-1.17	0.00	-0.58	0.25	-0.68***	1.08***	0.11	1.05*
	(0.29)	(0.17)	(1.40)	(0.99)	(0.73)	(0.71)	(2.78)	(1.77)	(5.36)	(1.04)	(0.52)	(0.52)	(0.09)	(0.09)	(0.26)	(0.62)
School type education option dummy	0.67***	0.44***	0.21	1.50	0.13	0.61***	2.48	0.52	0.00	0.58	1.50	0.51	-0.31***	0.72***	0.97***	0.49***
	(0.05)	(0.08)	(5.64)	(2.51)	(0.13)	(0.09)	(2.60)	(1.11)	(1.47)	(0.80)	(1.22)	(0.32)	(0.02)	(0.02)	(0.10)	(0.17)
School type screened dummy	1.19***	0.77***	1.43	1.38	0.06	0.73***	2.54	0.02	1.17	1.32	2.38*	0.01	0.01	0.94***	2.26***	0.00
	(0.04)	(0.05)	(1.82)	(1.17)	(0.19)	(0.19)	(2.09)	(2.69)	(1.76)	(0.99)	(1.24)	(0.50)	(0.03)	(0.02)	(0.10)	(0.19)
School type limited dummy	0.11	0.83***	0.48	2.82**	-0.55	0.78	0.87	2.38***	-1.42	0.73	-0.87	0.12	-1.80***	1.01***	-0.08	0.00
	(0.12)	(0.08)	(4.83)	(1.22)	(0.60)	(0.48)	(0.85)	(0.60)	(1.14)	(1.00)	(0.70)	(0.50)	(0.08)	(0.08)	(0.27)	(0.40)
School type audit dummy	-1.21***	2.31***	-1.42	0.08	-1.15	1.85***	-5.46***	7.20	-2.89***	2.12**	-0.56	0.04	-1.90***	1.85***	-3.44***	0.00
	(0.10)	(0.08)	(0.97)	(0.42)	(0.85)	(0.69)	(1.54)	(4.60)	(1.09)	(0.96)	(0.55)	(0.50)	(0.09)	(0.08)	(0.20)	(0.40)
High performance dummy (log-normal for $\beta a\,(1+fa)$)	-0.47***	1.28***	2.00*	1.31***	-2.99***	2.32***	0.08	0.16	8.41***	12.94***	0.87	0.63	-0.13***	1.30***	0.83*	0.20
	(0.07)	(0.11)	(1.11)	(0.43)	(0.81)	(0.86)	(0.48)	(0.65)	(1.54)	(0.39)	(1.87)	(0.51)	(0.05)	(0.10)	(0.50)	(0.41)
Tiny new school dummy	-3.79***	1.87***	0.13	0.25	-3.26***	2.35***	0.91	3.45***	-4.27***	2.05**	2.38***	1.75***	-3.05***	1.86***	3.33***	0.00
	(0.19)	(0.14)	(6.61)	(3.84)	(0.90)	(0.69)	(0.97)	(0.51)	(1.00)	(1.01)	(0.50)	(0.50)	(0.36)	(0.26)	(0.29)	(0.24)
Tiny old school dummy	-2.57***	0.00	-3.76***	0.00	-2.05***	0.42	-9.79*	5.66***	-3.49***	0.52	-9.98***	4.75***	-3.01***	0.81***	-6.75***	5.21***
	(0.13)	(0.90)	(1.21)	(0.63)	(0.42)	(1.22)	(5.26)	(1.48)	(0.83)	(1.55)	(0.61)	(0.64)	(0.04)	(0.05)	(0.15)	(0.52)
Small new school dummy	-1.68***	0.37	-1.47	0.01	-0.86*	0.01	0.34	0.04	-3.54**	0.01	-10.00***	5.66***	-1.44***	0.07	0.82	5.38***
	(0.10)	(0.27)	(3.92)	(0.95)	(0.51)	(0.44)	(4.35)	(5.70)	(1.69)	(1.01)	(0.50)	(0.48)	(0.09)	(0.51)	(0.54)	(0.42)
Small old school dummy	-1.53***	0.37***	-1.46	0.63	-1.06***	0.00	-3.15***	2.14***	-2.23***	0.00	-2.71***	0.72	-1.48***	0.40***	-0.69***	0.08
	(0.04)	(0.05)	(2.66)	(0.40)	(0.07)	(0.84)	(0.67)	(0.53)	(0.53)	(2.33)	(0.57)	(0.51)	(0.03)	(0.05)	(0.12)	(0.52)
Medium new school dummy	-0.74***	0.01	-1.93	0.07	-0.28	0.33	-1.08	0.97	-0.62	0.35	-4.68***	0.47	-0.49***	0.01	-0.37	0.18
	(0.10)	(0.85)	(2.05)	(0.95)	(0.54)	(0.55)	(1.08)	(0.66)	(2.31)	(0.45)	(0.57)	(0.50)	(0.16)	(0.86)	(0.35)	(0.29)
Medium old school dummy	-0.81***	0.00	-0.68	0.00	-0.50***	0.00	-0.47	1.26**	-1.48***	0.00	-0.99	0.00	-0.91***	0.42***	-0.39***	0.00
	(0.04)	(1.00)	(10.81)	(0.83)	(0.07)	(1.42)	(0.87)	(0.63)	(0.14)	(0.87)	(1.65)	(0.50)	(0.14)	(0.87)	(0.09)	(0.40)
N (students)	5,021				52,114				5,271				13,844			

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%

Notes: Based on the rational expectation model, this table shows the estimates of the mean and standard deviation of $(\beta_{ak}(1+f_{ak}))_k$ (the coefficient on X_{ask} in $t=0$) and $(\beta_{ak}f_{ak})_k$ (the coefficient in $t=1$). Standard errors are in parentheses. See Sections 3.1 and 3.3 for the details of the model and the estimation method, respectively. Appendix A.1 explains the construction of variables. See Section 3.4 for discussions about this table.

Table A.7: Estimates from the Rational Expectation Model: Reapplication Costs and Initial Assignment Effects

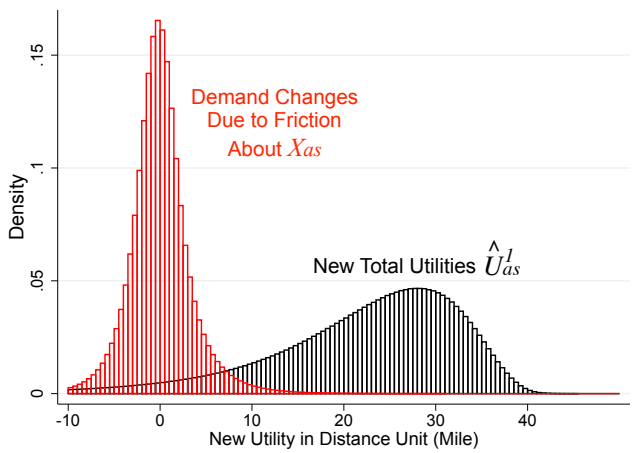
	Black or hispanic								White or asian							
	Grade 7 reading grade category: high				Low or middle				High				Low or middle			
	Reapplication cost *ca*		Initial assignment effect Y*a*		*ca*		Ya		*ca*		Ya		*ca*		*Ya*	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	1.11***	0.97	5.82*	3.26	4.65***	3.21**	6.68***	45.94***	1.41***	0.49	7.50***	2.23	3.16***	0.76	6.36***	2.59*
	(0.26)	(1.04)	(3.05)	(3.49)	(0.49)	(1.48)	(1.47)	(1.82)	(0.49)	(1.48)	(1.47)	(1.82)	(2.16)	(1.60)	(0.77)	(1.36)
N (students)	5,021				52,114				5,271				13,844			

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%

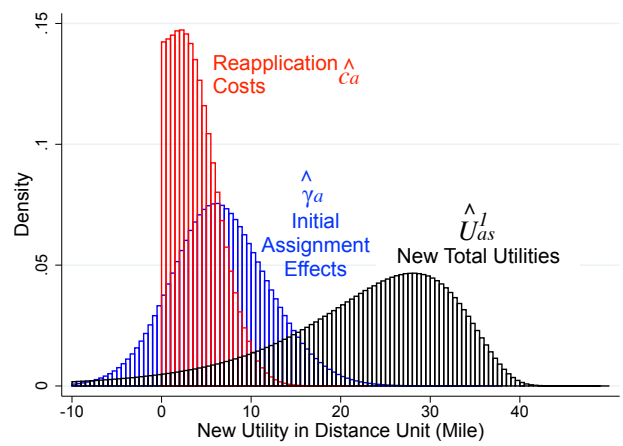
Notes: Based on the rational expectation model, this table shows the estimates of the mean and standard deviation of reapplication costs c_a and initial assignment effects γ_a. Standard errors are in parentheses. See Sections 3.1 and 3.3 for the details of the model and the estimation method, respectively. See Section 3.4 for discussions about this table.

Figure A.7: Summary of Estimates: Naive Free Expectation Model

(a) Demand Changes due to Learning



(b) Reapplication Costs & Initial Match Effects



Notes: Based on the naive free expectation model in Section 3.1, Panel 5a plots the distributions of estimated overall new utilities ($\hat{U}^1_{as}$) and latent demand changes due to frictions about observable school characteristics ($\Sigma_{k=1}^{K}\beta_{ak}f_{ak}X_{ask}$) for all (applicant a, school s) pairs. Panel 5b plots the distributions of estimated overall new utilities ($\hat{U}^1_{as}$), estimated reapplication costs ($\hat{c}_a$), and estimated initial assignment effects ($\hat{\gamma}_a$). On $\hat{c}_a$, it plots values implied by the estimated value of c_a/p_a and the rational expectation assumption that p_a is equal to the empirical probability of reapplication acceptance. Both panels are based on 50 simulations of the estimated model for each (applicant a, school s) pair. See Sections 3.1 and 3.3 for the details of the model and the estimation method, respectively. See Section 3.4 for discussions about this figure.

Table A.8: Estimates from the Naive Free Expectation Model: Preferences and Frictions

	Black or hispanic								White or asian							
	Grade 7 reading grade category: high				Low or middle				High				Low or middle			
	$\beta a(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$		$\beta a(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$		$\beta a(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$		$\beta a(1+fa)$ in $t=0$		Change $-\beta afa$ to $t=1$	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Negative distance (-10 miles, log-normal for $\beta a(1+fa)$)	0.90***	0.60***	3.34***	0.67	1.08***	0.59***	6.03***	2.29***	1.85***	0.74***	5.76***	2.60***	1.46***	0.66***	6.03***	2.29***
	(0.02)	(0.01)	(0.30)	(0.48)	(0.01)	(0.02)	(0.32)	(0.30)	(0.20)	(0.05)	(0.63)	(0.68)	(0.01)	(0.01)	(0.38)	(0.34)
School type unscreened dummy	-0.56*	1.27***	-2.69***	1.86*	-0.75	1.02	-0.99**	0.04	-1.17	0.00	-2.33*	1.55	-0.68***	1.08***	-0.99*	0.04
	(0.29)	(0.17)	(1.02)	(1.05)	(0.73)	(0.71)	(0.46)	(0.94)	(5.36)	(1.04)	(1.37)	(1.30)	(0.09)	(0.09)	(0.59)	(1.15)
School type education option dummy	0.67***	0.44***	-1.30***	0.82	0.13	0.61***	-1.26***	0.66***	0.00	0.58	-1.03***	0.72	-0.31***	0.72***	-1.26***	0.66**
	(0.05)	(0.08)	(0.36)	(0.93)	(0.13)	(0.09)	(0.17)	(0.23)	(1.47)	(0.80)	(0.29)	(0.47)	(0.02)	(0.02)	(0.19)	(0.30)
School type screened dummy	1.19***	0.77***	-0.12	0.49	0.06	0.73***	-0.63***	0.88***	1.17	1.32	0.00	0.08	0.01	0.94***	-0.63***	0.88**
	(0.04)	(0.05)	(0.23)	(1.71)	(0.19)	(0.19)	(0.17)	(0.26)	(1.76)	(0.99)	(0.24)	(0.99)	(0.03)	(0.02)	(0.20)	(0.39)
School type limited dummy	0.11	0.83***	-0.67	1.65**	-0.55	0.78	-1.66***	0.00	-1.42	0.73	-1.07	0.07	-1.80***	1.01***	-1.66***	0.00
	(0.12)	(0.08)	(0.56)	(0.65)	(0.60)	(0.48)	(0.33)	(0.30)	(1.14)	(1.00)	(0.70)	(1.00)	(0.08)	(0.08)	(0.51)	(1.00)
School type audit dummy	-1.21***	2.31***	-1.20**	0.13	-1.15	1.85***	-1.83***	0.06	-2.89***	2.12**	-0.26	0.05	-1.90***	1.85***	-1.83***	0.06
	(0.10)	(0.08)	(0.50)	(0.75)	(0.85)	(0.69)	(0.60)	(0.50)	(1.09)	(0.96)	(0.49)	(1.00)	(0.09)	(0.08)	(0.45)	(0.81)
High performance dummy (log-normal for $\beta a(1+fa)$)	-0.47***	1.28***	1.45***	0.73	-2.99***	2.32***	3.18***	3.25***	8.41***	12.94***	3.22**	1.26	-0.13***	1.30***	3.18***	3.25***
	(0.07)	(0.11)	(0.50)	(0.44)	(0.81)	(0.86)	(0.12)	(1.00)	(1.54)	(0.39)	(1.30)	(1.17)	(0.05)	(0.10)	(0.91)	(0.70)
Tiny new school dummy	-3.79***	1.87***	-0.52	0.47	-3.26***	2.35***	0.75**	1.03	-4.27***	2.05**	-0.03	0.06	-3.05***	1.86***	0.75	1.03**
	(0.19)	(0.14)	(0.55)	(1.14)	(0.90)	(0.69)	(0.29)	(1.00)	(1.00)	(1.01)	(0.69)	(1.00)	(0.36)	(0.26)	(0.57)	(0.42)
Tiny old school dummy	-2.57***	0.00	-3.21***	0.18	-2.05***	0.42	-2.77***	0.00	-3.49***	0.52	-5.14***	1.89***	-3.01***	0.81***	-2.77***	0.00
	(0.13)	(0.90)	(0.39)	(0.96)	(0.42)	(1.22)	(0.27)	(0.30)	(0.83)	(1.55)	(0.86)	(0.73)	(0.04)	(0.05)	(0.32)	(1.06)
Small new school dummy	-1.68***	0.37	-1.98***	0.15	-0.86*	0.01	-1.30***	1.84***	-3.54**	0.01	-8.85	4.73*	-1.44***	0.07	-1.30	1.84**
	(0.10)	(0.27)	(0.58)	(1.00)	(0.51)	(0.44)	(0.33)	(0.53)	(1.69)	(1.01)	(6.36)	(2.56)	(0.09)	(0.51)	(0.84)	(0.73)
Small old school dummy	-1.53***	0.37***	-1.98***	0.45	-1.06***	0.00	-1.56***	0.01	-2.23***	0.00	-2.31***	0.28	-1.48***	0.40***	-1.56***	0.01
	(0.04)	(0.05)	(0.29)	(0.40)	(0.07)	(0.84)	(0.16)	(0.99)	(0.53)	(2.33)	(0.31)	(0.97)	(0.03)	(0.05)	(0.24)	(0.98)
Medium new school dummy	-0.74***	0.01	-2.33***	0.14	-0.28	0.33	-2.22***	1.38**	-0.62	0.35	-2.37**	0.63	-0.49***	0.01	-2.22***	1.38**
	(0.10)	(0.85)	(0.57)	(0.99)	(0.54)	(0.55)	(0.36)	(0.58)	(2.31)	(0.45)	(1.08)	(1.09)	(0.16)	(0.86)	(0.83)	(0.61)
Medium old school dummy	-0.81***	0.00	-0.79***	0.10	-0.50***	0.00	-0.92***	0.03	-1.48***	0.00	-1.68***	0.24	-0.91***	0.42***	-0.92***	0.03
	(0.04)	(1.00)	(0.22)	(0.58)	(0.07)	(1.42)	(0.14)	(0.28)	(0.14)	(0.87)	(0.36)	(1.00)	(0.14)	(0.04)	(0.17)	(1.00)
N (students)	5,021				52,114				5,271				13,844			

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%

Notes: This table shows the estimates of the mean and standard deviation of $(\beta_{ak}(1+f_{ak}))_k$ (the coefficient on X_{ask} in $t=0$) and $(\beta_{ak}f_{ak})_k$ (the coefficient in $t=1$). Standard errors are in parentheses. See Sections 3.1 and 3.3 for the details of the model and the estimation method, respectively. Appendix A.1 explains the construction of variables. See Section 3.4 for discussions about this table.

Table A.9: Estimates from the Naive Free Expectation Model: Reapplication Costs and Initial Assignment Effects

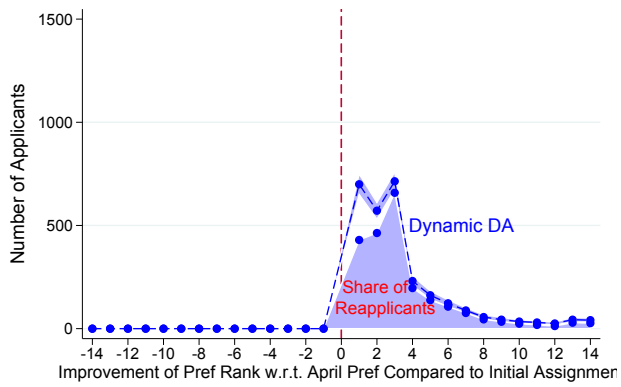
	Black or hispanic								White or asian							
	Grade 7 reading grade category: high				Low or middle				High				Low or middle			
	Scaled reapplication cost *ca/pa*		Initial assignment effect Y*a*		*ca/pa*		Ya		*ca/pa*		Ya		*ca/pa*		*Ya*	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	5.34***	5.44**	5.84***	3.74	15.90***	14.11***	6.02***	4.99***	19.12***	13.91***	6.79***	4.55***	15.90***	14.11***	6.02***	4.99***
	(1.07)	(2.31)	(0.74)	(2.58)	(0.87)	(1.81)	(1.47)	(1.21)	(2.66)	(3.17)	(1.45)	(1.28)	(1.60)	(1.13)	(0.67)	(1.26)
N (students)	5,021				52,114				5,271				13,844			

***: Significant at 1%, **: Significant at 5%, *: Significant at 10%

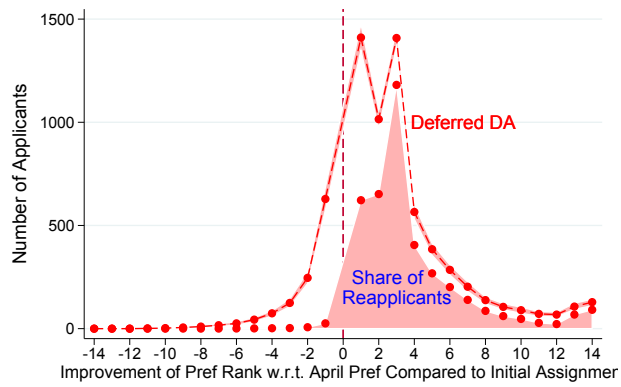
Notes: This table shows the estimates of the mean and standard deviation of scaled reapplication costs c_a/p_a and initial assignment effects γ_a. Standard errors are in parentheses. See Sections 3.1 and 3.3 for the details of the model and the estimation method, respectively. See Section 3.4 for discussions about this table.

Figure A.8: Centralized vs Discretionary Reapplication Processes (I): Details

(a) Dynamic DA Mechanism



(b) Deferred DA Mechanism



Notes: In Panel (a), the dotted line plots the distribution of the improvement of the preference rank of the finally assigned school under the dynamic deferred acceptance mechanism compared with the initial match. The preference rank is defined with respect to new preference $\succ_a^1$ defined in the main text. This distribution is conditional on applicants who get different assignments under the two mechanisms. The shaded area below the dotted line plots the same distribution as the dotted line conditional on applicants who reapply. Panel (b) does the same for the deferred deferred acceptance mechanism. The shaded area around each dotted line indicates the 95% simulation confident interval over 200 simulations of lottery numbers used by the mechanisms to break ties in priorities.