NBER WORKING PAPER SERIES

RESEARCH DESIGN MEETS MARKET DESIGN: USING CENTRALIZED ASSIGNMENT FOR IMPACT EVALUATION

Atila Abdulkadiroglu
Joshua D. Angrist
Yusuke Narita
Parag A. Pathak

Working Paper 21705
http://www.nber.org/papers/w21705

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2015

We thank Alisha Chiarelli, Brian Eschbacher, Van Schoales, and the staff at Denver Public Schools for answering our questions and facilitating access to data. Isaiah Andrews, Eduardo Azevedo, Gary Chamberlain, Victor Chernozhukov, Jerry Hausman, Peter Hull, Hide Ichimura, Guido Imbens, Rafael Lalive, Edward Lazear, Jacob Leshno, Anna Mikuscheva, Chris Walters and seminar participants at Harvard, SOLE, Stanford University, the Fall 2015 NBER market design meeting, Kyoto, Osaka, and Tokyo provided helpful feedback. We're especially indebted to Mayara Felix and Ye Ji Kee for expert research assistance and to MIT SEII program manager Annice Correia for invaluable administrative support. We gratefully acknowledge funding from the Laura and John Arnold Foundation and the National Science Foundation (under awards SES-1056325 and SES-1426541). Data from Denver Public Schools were made available to us through the Institute for Innovation in Public School Choice. Abdulkadiroglu and Pathak are Scientific Advisory Board members of the Institute for Innovation in Public School Choice. Angrist's daughter teaches at a Boston charter school. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w21705.ack

Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation
Atila Abdulkadiroglu, Joshua D. Angrist, Yusuke Narita, and Parag A. Pathak
NBER Working Paper No. 21705
November 2015
JEL No. C26,D47,I20

ABSTRACT

A growing number of school districts use centralized assignment mechanisms to allocate school seats in a manner that reflects student preferences and school priorities. Many of these assignment schemes use lotteries to ration seats when schools are oversubscribed. The resulting random assignment opens the door to credible quasi-experimental research designs for the evaluation of school effectiveness. Yet the question of how best to separate the lottery-generated variation integral to such designs from non-random preferences and priorities remains open. This paper develops easily-implemented empirical strategies that fully exploit the random assignment embedded in the widely-used deferred acceptance mechanism and its variants. We use these methods to evaluate charter schools in Denver, one of a growing number of districts that integrate charter and traditional public schools in a unified assignment system. The resulting estimates show large achievement gains from charter school attendance. Our approach expands the scope for impact evaluation by maximizing the number of students and schools that can be studied using random assignment. We also show how to use DA to identify causal effects in models with multiple school sectors.

Atila Abdulkadiroglu
Duke University
Department of Economics
Durham, NC 27708
atila.abdulkadiroglu@duke.edu

Yusuke Narita
Department of Economics, MIT
77 Massachusetts Avenue
Cambridge, MA 02139
narita@mit.edu

Joshua D. Angrist
Department of Economics, E17-226
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
angrist@mit.edu

Parag A. Pathak
Department of Economics, E17-240
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
ppathak@mit.edu

A data appendix is available at:
http://www.nber.org/data-appendix/w21705

1 Introduction

Families in many large urban districts can now apply for seats at any public school in their district. The fact that some schools are more popular than others and the need to distinguish between students who have different priorities at a given school generates a matching problem. Introduced by Gale and Shapley (1962) and Shapley and Scarf (1974), matchmaking via market design allocates scarce resources, such as seats in public schools, in markets where prices cannot be called upon to perform this function. The market-design approach to school choice, pioneered by Abdulkadiroğlu and Sönmez (2003), is used in a long and growing list of public school districts in America, Europe, and Asia. Most of these cities match students to schools using a mechanism known as deferred acceptance (DA).

Two benefits of matchmaking schemes like DA are efficiency and fairness: the resulting match improves welfare and transparency relative to ad hoc alternatives, while lotteries ensure that students with the same preferences and priorities have the same chance of obtaining highly-sought-after seats. DA and related algorithms also have the virtue of narrowing the scope for strategic behavior that would otherwise give sophisticated families the opportunity to manipulate an assignment system at the expense of less-sophisticated participants (Abdulkadiroğlu et al., 2006; Pathak and Sönmez, 2008). No less important than these economic considerations is the fact that centralized assignment generates valuable data for empirical research on schools. In particular, when schools are oversubscribed, lottery-based rationing generates quasi-experimental variation in school assignment that can be used for credible evaluation of individual schools and of school reform models like charters.

Previous research exploiting the lotteries embedded in DA include studies of schools in Charlotte-Mecklenburg (Hastings et al., 2009; Deming, 2011; Deming et al., 2014) and New York (Bloom and Unterman, 2014; Abdulkadiroğlu et al., 2013). Causal effects in these studies are identified by compelling quasi-experimental variation, but the research designs deployed in this work fail to exploit the full power of the random assignment embedded in centralized assignment schemes. A major stumbling block in this context is the elaborate multi-stage nature of market-design matching. Market design weaves random assignment into an elaborate tapestry of information on student preferences and school priorities. In principle, all features of student preferences and school priorities can shape the probability of assignment to each school. Preferences and priorities are far from randomly assigned, of course. Families tend to prefer schools located in their neighborhoods, for example, while schools may grant priority to children poor enough to qualify for a subsidized lunch. It's only conditional on preferences and priorities, therefore, that DA-generated assignments are independent of potential outcomes.

To eliminate the selection bias that arises from the dependence of assignments on preferences and priorities, research exploiting centralized assignment has focused either on offers of seats at students' first choice schools, or relied on instrumental variables (IVs) indicating whether a student's lottery number falls below the highest number offered a seat at all schools he's ranked (we call this a qualification instrument). The first choice strategy conditions on the identity of the school ranked first, while qualification instruments condition on the set of schools ranked. These IV strategies are likely to produce estimates free of omitted variables bias. At the same

time, both first-choice and qualification instruments discard much of the variation induced by DA.

This paper explains how to recover the full range of quasi-experimental variation embedded in centralized assignment. Specifically, we show how DA maps information on preferences, priorities, and school capacities into a conditional probability of random assignment, often referred to as the propensity score. As in other stratified randomized research designs, conditioning on the propensity score eliminates selection bias arising from the association between all conditioning variables and potential outcomes (Rosenbaum and Rubin, 1983). The payoff to propensity-score conditioning turns out to be substantial in our application: full stratification on preferences and priorities reduces degrees of freedom markedly, eliminating many schools and students from consideration, while score-based stratification leaves our research sample largely intact. The propensity score does more for us than reduce dimensionality, however. Because all applicants with score values strictly between zero and one contribute variation that can be used for evaluation, the propensity score identifies the maximal set of applicants for whom we have a randomized school-assignment experiment.

The propensity score generated by DA-type mechanisms does not have a general closed form solution. Our theoretical framework therefore revolves around an asymptotic "large market" approximation to the finite-market score. This *DA propensity score* is a function of a few easily-computed sample statistics. We also construct propensity score estimates using simulation, that is, by drawing lottery numbers many times and computing the resulting average assignment rates across draws. Both the simulated and DA (analytic) propensity scores work well as far as covariate balance goes, but the approximate formula is, of course, much more quickly computed, and highlights specific sources of randomness and confounding in DA-based assignment schemes. In other words, the DA propensity score reveals the nature of the stratified experimental design embedded in a particular match.

Our test bed for the DA propensity score is an empirical analysis of charter school effects in the Denver Public School (DPS) district, a new and interesting setting for charter school impact evaluation.[1] Because DPS assigns seats at traditional and charter schools in a unified match, the population attending DPS charters is less positively selected than in large urban districts with decentralized charter lotteries. This context makes DPS charter effects relevant for the ongoing debate over charter expansion. As far as we know, ours is the first charter evaluation to exploit an assignment scheme that simultaneously allocates seats in both the charter and traditional public school sectors.

The next section uses simple market design examples to explain the problem at hand. Following this explanation, Section 3 uses the theory of market design to characterize the propensity

[1]Charter schools operate with considerably more independence than traditional public schools. Among other differences, many charters fit more instructional hours into a year by running longer school days and providing instruction on weekends and during the summer. Because few charter schools are unionized, they hire and fire teachers and administrative staff without regard to the collectively bargained seniority and tenure provisions that constrain such decisions in many public schools. About half of Denver charters implement versions of the *No Excuses* model of urban education. No Excuses charters run a long school day and year, emphasize discipline and comportment and traditional reading and math skills, and rely heavily on data and teacher feedback to improve instruction. For more background on the charter sector, see Abdulkadiroğlu et al. (2011) and Angrist et al. (2013).

score for DA offers in large markets. Section 4 applies these results to estimate charter effects. Our empirical evaluation strategy uses an indicator for DA-generated charter offers as an instrument for charter school attendance in a two-stage least squares (2SLS) setup. This 2SLS procedure eliminates bias from non-random variation in preferences and priorities by controlling for the DA propensity score. This section also shows how to estimate effects for multiple sectors. Specifically, we look at DPS innovation schools, a popular alternative to the charter model. Finally, Section 5 summarizes our theoretical and empirical findings and outlines an agenda for further work. A theoretical appendix derives propensity scores for the Boston (Immediate Acceptance) mechanism and for DA with multiple tie-breaking.

2 Understanding the DA Propensity Score

We begin by reviewing the basic DA setup for school choice, showing how DA generates probabilities of school assignment that depend on preferences, priorities, and capacities.

A total of n students are to be assigned seats at schools of varying capacities. Students report their preferences by ranking schools on an application form or website, while schools rank students by placing them in priority groups. For example, a school may give the highest priority to students with already-enrolled siblings, second highest priority to those who live nearby, with the rest in a third priority group below these two. Each student is also randomly assigned a lottery number that distinguishes between those with the same preferences and priorities. DA assigns students to schools like this:

> Each student applies to his most preferred school. Each school ranks all its applicants first by priority then by random number within priority groups and *tentatively* admits the highest-ranked applicants in this order up to its capacity. Other applicants are rejected.
>
> Each rejected student applies to his next most preferred school. Each school ranks these new applicants *together with applicants that it admitted tentatively in the previous round*, first by priority and then by random number. From this pool, the school tentatively admits those it ranks highest up to capacity, rejecting the rest.

This algorithm terminates when there are no new applications (some students may remain unassigned).

DA produces a stable allocation in the following sense: any student who prefers another school to the one he has been assigned must be outranked at that school, either because everyone assigned there has higher priority, or because those who share the student's priority at that school have higher lottery numbers. DA is also strategy-proof, meaning that families do as well as possible by submitting a truthful preference list (for example, there is nothing to be gained by ranking under-subscribed schools highly just because they are likely to yield seats). See Roth and Sotomayor (1990) for a review of these and related theoretical results.

2.1 Propensity Score Pooling

Within a given market structure (defined by the total number of applicants and the number of seats to be filled at each school), the probability that DA assigns student i a seat at school s depends on student i's preferences and priority status at s and elsewhere. We refer to a student's preferences and priorities as student *type*. For example, a student of one type might rank school b first, school a second, and have sibling priority at b.

Suppose we'd like to estimate the causal effect of attending a particular school relative to other schools that students who rank this school might attend (our empirical work focuses on the causal effect of attendance at groups of schools, but the logic behind such comparisons is similar). Families exhibit a strong preference for schools in their neighborhood, so offer rates are high for nearby schools. At the same time, there are important differences in average achievement across Denver neighborhoods. Consequently, families who live in lower-income neighborhoods are more likely to be offered seats in schools attended by low achievers for reasons unrelated to the causal effects of school attendance. This is important in our context because there are more charters in poor areas. Likewise, selection bias can arise from failure to control for priorities: in some districts, for example, students poor enough to qualify for a subsidized lunch are granted priority for seats at schools of one sort or another.

DA treats students of the same type symmetrically in that everyone of a given type faces the same probability of assignment to each school. In other words, conditional on type, all that remains to determine school assignment is a random number that is independent of student characteristics and potential outcomes (this intuitive claim is proved in the next section). We can therefore eliminate selection bias in comparisons of those who are and aren't offered seats at particular schools simply by conditioning on type. As a practical matter, however, we'd like to avoid full type conditioning, since this produces many small and even singleton or empty cells, reducing the sample available for impact analysis dramatically. The following simple example illustrates this point.

Example 1. Five students $\{1,2,3,4,5\}$ apply to three schools $\{a,b,c\}$, each with one seat. Student 5 has the highest priority at c and student 2 has the highest priority at b, otherwise the students have the same priority at all schools. We're interested in measuring the effect of an offer at school a. Student preferences are

$$\begin{aligned} 1 &: a \succ b, \\ 2 &: a \succ b, \\ 3 &: a, \\ 4 &: c \succ a, \\ 5 &: c, \end{aligned}$$

where $a \succ b$ means that a is preferred to b. Students 3 and 5 find only a single school acceptable.

Note that no two students here have the same preferences and priorities. Consequently, full-type stratification puts each student into a different stratum. This rules out research strategies that rely on full type conditioning to eliminate selection bias. But full type conditioning is unnecessary in this case because DA assigns students 1, 2, 3, and 4 to school a each with

probability 0.25. This calculation reflects the fact that 5 beats 4 at c by virtue of his priority there, leaving 1, 2, 3, and 4 all applying to a in the second round and no one advantaged there. The impact of assignment to a can therefore be analyzed in a single stratum containing four students. As we show formally in Section 3, this stratification scheme is determined by the propensity score, the conditional probability of random assignment to a. Specifically, we can use a dummy indicating offers at a as an instrument for attendance at a in a sample that includes types 1-4: offers among these types occur with equal frequency and are therefore independent of potential outcomes and student characteristics. At the same time, the fact that offers of seats at school a boost enrollment there generates the relevant first stage.

2.2 Further Pooling in Large Markets

Under DA, the propensity score for assignment to school a is determined both by a student's failure to win a seat at schools he ranks more highly than a and by the odds he wins a seat at a in competition with those who have also ranked a and similarly failed to find seats at schools they've ranked more highly. This two-part structure leads to a large-market approximation that generates pooling beyond that provided by the finite-market propensity score. We illustrate this point via a second simple example.

Example 2. Four students $\{1, 2, 3, 4\}$ apply to three schools $\{a, b, c\}$, each with one seat. There are no school priorities and student preferences are

$$\begin{aligned} &1 : c, \\ &2 : c \succ b \succ a, \\ &3 : b \succ a, \\ &4 : a. \end{aligned}$$

As in Example 1, each student is of a different type.

Let $p_a(i)$ for $i = 1, 2, 3, 4$ denote the probability that type i is assigned school a. With four students, $p_a(i)$ comes from $4! = 24$ possible lottery draws, all equally likely. Given this modest number of possibilities, $p_a(i)$ is easily calculated by enumeration:

- Not having ranked a, type 1 is never assigned there, so $p_a(1) = 0$.
- Type 2 is seated at a when schools he's ranked ahead of a, schools b and c, are filled by others, and when he also beats type 4 in competition for a seat at a. This occurs for the two realizations of the form $(s, t, 2, 4)$ for $s, t = 1, 3$. Therefore, $p_a(2) = 2/24 = 1/12$.
- Type 3 is seated at a when the schools he's ranked ahead of a–in this case, only b–are filled by others, while he also beats type 4 in competition for a seat at a. b can be filled by type 2 only when 2 loses to 1 in the lottery at c. Consequently, type 3 is seated at a only in a sequence of the form $(1, 2, 3, 4)$, which occurs only once. Therefore, $p_a(3) = 1/24$.
- Finally, since type 4 gets the seat at a if and only if the seat does not go to type 2 or type 3, $p_a(4) = 21/24$.

In this example, the propensity score differs for each student. But in larger markets with the same distribution of types, the score is smoother. To see this, consider a large market that replicates the structure of this example n times, so that n students of each type apply to 3 schools, each with n seats.[2] With large n, enumeration of assignment possibilities is a chore. We can, however, simulate the propensity score by repeatedly drawing lottery numbers.

The relationship between simulated probabilities of assignment and market size for Example 2, plotted in Figure 1, reveals that as the market grows, the distinction between types 2 and 3 disappears. In particular, Figure 1 shows that for large enough n,

$$p_a(2) = p_a(3) = 1/12; \quad p_a(1) = 0; \quad p_a(4) = 10/12 = 5/6,$$

with the probability of assignment at a for types 2 and 3 converging quickly. This convergence is a consequence of a result established in the next section, which shows that the large-market probabilities that types and 2 and 3 are seated at a are both determined by failure to win a seat at b. The fact that student 3 ranks c ahead of b is irrelevant.

A patient analyst can always approximate a finite market score by simulation, but our large-market results reveal why some schools and applicant types are subject to random assignment (even at schools that are under-subscribed), why applicants of different types share the same risk, and why for some applicants, assignment risk is degenerate (even at schools that are over-subscribed). A signal feature of the large market characterization is the role played by lottery qualification cutoffs at schools ranked ahead of school a in determining probabilities of assignment at a. This is illustrated by Example 2, which shows that, in the large-market limit, among schools that an applicant prefers to a, we need only be concerned with what happens at the school at which it's easiest to qualify. In general, this *most informative disqualification* (MID) determines how distributions of lottery numbers for applicants of differing types are effectively truncated before entering the competition for seats at a. As we show below, the fact that the large market score depends on type only through a set of constructs like MID allows us to replace full type conditioning with something much smoother.

3 Score Theory

3.1 Setup

A general school choice problem, which we refer to as an economy, is defined by a set of students, schools, school capacities, student preferences over schools, and student priorities at schools. Let I denote a set of students, indexed by i, and let $s = 1, ..., S$ index schools. We consider markets with a finite number of schools, but with either finite (n) or infinitely many students. As in Abdulkadiroğlu et al. (2015) and Azevedo and Leshno (2014), the latter setting is referred to as a *continuum economy*. In a continuum economy, $I = [0, 1]$ and school capacities are defined as the fraction of the continuum that can be seated at each school.

[2] Many market-design analysts have found this sort of large-market approximation useful. Examples include Abdulkadiroğlu et al. (2015); Azevedo and Leshno (2014); Budish (2011); Che and Kojima (2010); Kesten and Ünver (2015).

Student i's preferences over schools constitute a partial ordering of schools, $\succ_i$, where $a \succ_i b$ means that i prefers school a to school b. Each student is also granted a priority at every school. Let $\rho_{is} \in \{1, ..., K, \infty\}$ denote student i's priority at school s, where $\rho_{is} < \rho_{js}$ means school s prioritizes i over j. For instance, $\rho_{is} = 1$ might encode the fact that student s has sibling priority at school s, while $\rho_{is} = 2$ encodes neighborhood priority, and $\rho_{is} = 3$ for everyone else. We use $\rho_{is} = \infty$ to indicate that i is ineligible for school s. Many students share priorities at a given school, in which case $\rho_{is} = \rho_{js}$ for some $i \neq j$. Let $\boldsymbol{\rho}_i = (\rho_{i1}, ..., \rho_{iS})$ be the vector of student i's priorities for each school. *Student type* is denoted by $\theta_i = (\succ_i, \boldsymbol{\rho}_i)$. We say that a student of type θ has preferences $\succ_\theta$ and priorities ρ_θ. Θ denotes the set of all possible types.

An economy is also characterized in part by a non-negative capacity vector, $\mathbf{q}$, which is normalized by the total number of students, or by their measure when students are indexed continuously. In a finite economy, where the set I contains n students and each school s has k_s seats, capacity is defined by $q_s = \frac{k_s}{n}$. In a continuum economy, q_s is the proportion of the set I that can be seated at school s.

The analysis here is concerned with school assignment mechanisms that use lotteries to distinguish between students with the same preferences and priorities. Student i's lottery number, r_i, is the realization of a uniformly distributed random variable on $[0, 1]$, independent and identically distributed for all students. In particular, lottery draws are independent of type. In what follows, we consider a centralized assignment system relying on a single lottery number for each student. Extension to the less-common multiple tie-breaking case, in which a student may have different lottery numbers at different schools, is discussed in the theoretical appendix.

For any set of student types $\Theta_0 \subset \Theta$ and for any number $r_0 \in [0, 1]$, define the set of students in Θ_0 with lottery number less than r_0 to be

$$I(\Theta_0, r_0) = \{i \in I \mid \theta_i \in \Theta_0, r_i \leq r_0\}.$$

We use the shorthand notation $I_0 = I(\Theta_0, r_0)$ for sets of applicants defined by type and lottery number. Also, when $r_0 = 1$, so that I_0 includes all lottery numbers, the second argument is omitted and $I_0 = \{i \in I \mid \theta_i \in \Theta_0\}$ for various choices of Θ_0.

When discussing a continuum economy, we let $F(I_0)$ denote the fraction of students in I_0. Since lottery numbers are uniform and independent of type, this is given by

$$F(I_0) = E[1\{\theta_i \in \Theta_0\}] \times r_0,$$

where $E[1\{\theta_i \in \Theta_0\}]$ is the proportion of types in set Θ_0. In a finite economy with n students, the corresponding fraction is computed as

$$F(I_0) = \frac{|I_0|}{n}.$$

For a continuum economy, $F(I_0)$ is fixed, that is, non-stochastic. By contrast, $F(I_0)$ for a finite economy depends on the realized lottery draw. Either way, the student side of an economy is fully characterized by the distribution of types and lottery numbers, for which we sometimes use the shorthand notation, F. Note also that every finite economy has a continuum analog. This analog can be constructed by replicating the type distribution and the number of seats in the finite economy, while fixing the proportion of seats at school s to be q_s.

Defining DA

We define DA using the notation outlined above, nesting the finite-market and continuum cases. First, combine priority status and lottery realization into a single number for each student and school, the *student rank*:

$$\pi_{is} = \rho_{is} + r_i.$$

Since the difference between any two priorities is at least 1 and random numbers are between 0 and 1, student rank is lexicographic in priority and lottery numbers.

DA proceeds in a series of rounds. Denote the evolving vector of *admissions cutoffs* in round t by $\mathbf{c}^t = (c_1^t, ..., c_S^t)$. The *demand* for seats at school s conditional on $\mathbf{c}^t$ is defined as

$$Q_s(\mathbf{c}^t) = \{i \in I \mid \pi_{is} \leq c_s^t \text{ and } s \succ_i \tilde{s} \text{ for all } \tilde{s} \in S \text{ such that } \pi_{i\tilde{s}} \leq c_{\tilde{s}}^t\}.$$

In other words, school s is demanded by students with rank below the school-s cutoff, who prefer school s to any other school for which they are also below the relevant cutoff.

The largest possible value of an eligible student's rank is $K+1$, so we can start with $c_s^1 = K+1$ for all s. Cutoffs then evolve as follows:

$$c_s^{t+1} = \begin{cases} K+1 & \text{if } F(Q_s(\mathbf{c}^t)) < q_s, \\ \max\left\{x \in [0, K+1] \mid F(\{i \in Q_s(\mathbf{c}^t) \text{ such that } \pi_{is} \leq x\}) \leq q_s\right\} & \text{otherwise;} \end{cases}$$

where, because the argument for F can be written in the form $\{i \in I \mid \theta_i \in \Theta_0, r_i \leq r_0\}$, the expression is well-defined. This formalizes the idea that when the demand for seats at s falls below capacity at s, the cutoff is $K+1$. Otherwise, the cutoff at s is the largest value such that demand for seats at s is less than or equal to capacity at s.

The final admissions cutoffs determined by DA for each school s are given by

$$c_s = \lim_{t \to \infty} c_s^t.$$

The set of students that are assigned school s under DA is the demand for seats at the limiting cutoffs: $\{i \in Q_s(\mathbf{c})\}$ where $\mathbf{c} = (c_1, ..., c_S)$. Since $c_s \leq K+1$, an ineligible student is never assigned to school s.

We write the final DA cutoffs as a limiting outcome to accommodate the continuum economy; in finite markets, DA converges in a finite number of rounds. Appendix A.1 shows that this description of DA is valid in the sense that: (a) the necessary limits exist for every economy, finite or continuous; (b) for every finite economy, the allocation upon convergence matches that produced by DA as usually described (for example, by Gale and Shapley (1962) and the many market design studies building on their work).

Defining the Propensity Score

DA-generated offers depend on preferences and priorities as well as on lottery numbers. DA can therefore be seen as inducing a stratified randomized trial, where the "strata" are defined by type, θ. Because students of different types are likely to have different outcomes for reasons unrelated to their DA assignments, we're interested in isolating the variation in offers determined by lottery numbers alone.

As in Rosenbaum and Rubin (1983)'s classic analysis of covariate conditioning, a key component of our effort to isolate random assignment within strata is the *propensity score*. The propensity score for a market of any size, denoted $p_s(\theta)$, is the scalar function of type defined by

$$p_s(\theta) = \Pr[D_i(s) = 1|\theta_i = \theta],$$

where $D_i(s)$ indicates whether student i is offered a seat at school s. This function has domain given by the set of types who rank s. We think of this as the group of applicants to s; for the moment, the notation $p_s(\theta)$ ignores the fact that the propensity score depends on market size.

Propensity score conditioning is motivated by a pair of conditional independence results. We first have the fact that DA offers are randomly assigned conditional on student type. In other words, for any random variable W_i that is independent of lottery numbers (this can be anything that is not a function of lottery numbers, including potential outcomes and student characteristics like free lunch status), the offer distribution satisfies

$$P[D_i(s) = 1|W_i, \theta_i = \theta] = P[D_i(s) = 1|\theta_i = \theta]. \tag{1}$$

Although unsurprising, this result provides a necessary foundation for everything that follows; Appendix A.2 therefore presents a formal proof.

The examples in Section 2 show that full type conditioning, that is, conditioning on each value of θ, reduces the sample available for impact evaluation and can eliminate schools and students from a causal analysis. It's natural, therefore, to consider grouping and smoothing schemes that implicitly pool values of θ. The propensity score theorem tells us how this pooling can be accomplished while still ensuring against omitted variables bias from any association between student type and potential outcomes. Given (1), it follows from Rosenbaum and Rubin's (1983) propensity score theorem (and from our proof of (1)) that propensity score conditioning is enough to ensure that offers are independent of W_i. In other words,

$$P[D_i(s) = 1|W_i, p_s(\theta_i) = p_s(\theta)] = P[D_i(s) = 1|p_s(\theta_i) = p_s(\theta)] = p_s(\theta). \tag{2}$$

Equation (2) implies that propensity score conditioning eliminates the possibility of omitted variables bias due to the dependence of offers on type.[3]

At first blush the conditional independence property described by Equation (2) might seem to be of little practical value: knowing nothing about the functional form of $p_s(\theta)$, we are left with as many possible score values as there are types. Our next step, therefore, is to derive an expression for $p_s(\theta)$ that exploits the structure of DA, showing how the score generated by DA indeed pools types.

3.2 Characterizing the DA Propensity Score

A key component in our characterization of $p_s(\theta)$ is the notion of a *marginal priority* group at school s. The marginal priority group consists of applicants for whom seats are allocated by

[3]Rosenbaum and Rubin (1983) also show that the propensity score is the coarsest balancing score, which in this case means that no coarser function of type ensures conditional independence of $D_i(s)$ and W_i. Hahn (1998), Hirano et al. (2003), and Angrist and Hahn (2004) discuss the efficiency consequences of conditioning on the score.

lottery if the school is over-subscribed. Formally, the marginal priority, ρ_s, is the integer part of the cutoff, c_s. Conditional on being rejected by all more preferred schools and applying for school s, a student is assigned s with certainty if his $\rho_{is} < \rho_s$, that is, if he clears marginal priority. Applicants with $\rho_{is} > \rho_s$ have no chance of finding a seat at s. Applicants for whom $\rho_{is} = \rho_s$ are marginal: these applicants are seated at s when their lottery numbers fall below a school-specific *lottery cutoff*. The lottery cutoff at school s, denoted τ_s, is the decimal part of the cutoff at s, that is, $\tau_s = c_s - \rho_s$.

These observations motivate a partition determined by marginal priorities at s. Let Θ_s denote the set of student types who rank s and partition Θ_s according to

i) $\Theta_s^n = \{\theta \in \Theta_s \mid \rho_{\theta s} > \rho_s\}$, (*never seated*)

ii) $\Theta_s^a = \{\theta \in \Theta_s \mid \rho_{\theta s} < \rho_s\}$, (*always seated*)

iii) $\Theta_s^c = \{\theta \in \Theta_s \mid \rho_{\theta s} = \rho_s\}$. (*conditionally seated*)

The set Θ_s^n contains applicant types who have worse-than-marginal priority at s. No one in this group is assigned to s. Θ_s^a contains applicant types that clear marginal priority at s. Some of these applicants may end up seated at a school they prefer to s, but they're assigned s for sure if they fail to find a seat at any school they've ranked more highly. Finally, Θ_s^c is the subset of Θ_s that is marginal at s. These applicants are assigned s when they're not assigned a higher choice *and* have a lottery number that clears the lottery cutoff at s.

A second key component of our score formulation reflects the fact that failure to qualify at schools other than s may truncate the distribution of lottery numbers in the marginal priority group for s. To characterize the distribution of lottery numbers among those at risk of assignment at s, we first define the set of schools ranked above s. Specifically, applicants of type θ view the following set of schools as better than s:

$$B_{\theta s} = \{s' \in S \mid s' \succ_\theta s\}.$$

An applicant's *most informative disqualification* (MID) at s is defined as a function of the cutoffs at schools in $B_{\theta s}$

$$MID_{\theta s} \equiv \begin{cases} 0 & \text{if } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s}, \\ 1 & \text{if } \rho_{\theta\tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \\ \max\{\tau_{\tilde{s}} \mid \tilde{s} \in B_{\theta s} \text{ and } \rho_{\theta\tilde{s}} = \rho_{\tilde{s}}\} & \text{if } \rho_{\theta\tilde{s}} = \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s} \text{ and } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ otherwise.} \end{cases}$$

$MID_{\theta s}$ tells us how the lottery number distribution among applicants to s is truncated by qualification at schools these applicants prefer to s. $MID_{\theta s}$ is zero when type θ students have worse-than-marginal priority at all higher ranked schools: when no s applicants can be seated at a more preferred school, there's no lottery number truncation among those at risk of assignment to s. On the other hand, when at least one school in $B_{\theta s}$ is under-subscribed, no one of type θ competes for a seat at s. Truncation is therefore complete, and $MID_{\theta s} = 1$.

The definition of $MID_{\theta s}$ also reflects the fact that, among applicants for whom $\rho_{\theta\tilde{s}} = \rho_{\tilde{s}}$ for some $\tilde{s} \in B_{\theta s}$, any student who fails to clear $\tau_{\tilde{s}}$ is surely disqualified at schools with lower cutoffs.

For example, applicants who fail to qualify at a school with a cutoff of 0.5 fail to qualify at schools with cutoffs below 0.5. Therefore, to keep track of the truncation induced by disqualification at all schools an applicant prefers to s, we need to record only the most forgiving cutoff that an applicant fails to clear.

In finite markets, $MID_{\theta s}$ varies from one lottery draw to another, but in a continuum economy, $MID_{\theta s}$ is fixed. Consider the large-market analog of Example 2 in which n students of each of four types compete for the n seats at each of three schools. In this example, there's a single priority group, so everyone is marginal. For large n, we can think of realized lottery numbers as being distributed according to a continuous uniform distribution over $[0, 1]$. Types 2 and 3 rank different schools ahead of a, i.e., $B_{3a} = \{b\}$ while $B_{2a} = \{b, c\}$. Nevertheless, because $\tau_c = 0.5 < 0.75 = \tau_b$, we have that $MID_{2a} = MID_{3a} = \tau_b = 0.75$. To see where these cutoffs come from, note first that among the $2n$ type 1 and type 2 students who rank c first in this large market, those with lottery numbers lower (better) than 0.5 are assigned to c since it has a capacity of n: $\tau_c = 0.5$. The remaining type 2 students (half of the original mass of type 2), all of whom have lottery numbers higher (worse) than 0.5, must compete with all type 3 students for seats at b. We therefore have $1.5n$ school-b hopefuls but only n seats at b. All type 3 students with lottery numbers below 0.5 get seated at b (the type 2 students all have lottery numbers above 0.5), but this doesn't fill b. The remaining seats are therefore split equally between type 2 and 3 students in the upper half of the lottery distribution, implying that the highest lottery number seated at b is $\tau_b = 0.75$.

The following theorem uses the marginal priority and MID concepts to define an easily-computed *DA propensity score* that is a deterministic function of applicant type:

Theorem 1. *Consider a continuum economy populated by applicants of type $\theta \in \Theta$ to be assigned to schools indexed by $s \in S$. For all s and θ in this economy, we have:*

$$p_s(\theta) = \varphi_s(\theta) \equiv \begin{cases} 0 & \text{if } \theta \in \Theta_s^n, \\ (1 - MID_{\theta s}) & \text{if } \theta \in \Theta_s^a, \\ (1 - MID_{\theta s}) \times \max\left\{0, \dfrac{\tau_s - MID_{\theta s}}{1 - MID_{\theta s}}\right\} & \text{if } \theta \in \Theta_s^c. \end{cases} \tag{3}$$

where we also set $\varphi_s(\theta) = 0$ when $MID_{\theta s} = 1$ and $\theta \in \Theta_s^c$.

The proof appears in Appendix A.3.

The case without priorities offers a revealing simplification of this result. Without priorities, DA is the same as a random serial dictatorship (RSD), that is, a serial dictatorship with applicants ordered by lottery number (see, e.g., Abdulkadiroglu and Sonmez 1998, Svensson 1999, Pathak and Sethuraman 2010).[4] Theorem 1 therefore implies the following corollary, which gives the *RSD propensity score*:

[4]Exam school seats are often assigned by a serial dictatorship based on admission test scores instead of random numbers (see, e.g., Abdulkadiroğlu et al. 2014a, Dobbie and Fryer 2014). A generalization of RSD, multi-category serial dictatorship, is used for Turkish college admissions (Balinski and Sönmez, 1999).

Corollary 1. *Consider a continuum economy with no priorities populated by applicants of type $\theta \in \Theta$, to be assigned to schools indexed by $s \in S$. For all s and θ in this economy, we have:*

$$\varphi_s(\theta) \equiv (1 - MID_{\theta s}) \times \max\left\{0, \frac{\tau_s - MID_{\theta s}}{1 - MID_{\theta s}}\right\} = \max\left\{0, \tau_s - MID_{\theta s}\right\}.$$

Without priorities, Θ_s^n and Θ_s^a are empty. The probability of assignment at s is therefore determined solely by draws from the truncated distribution of lottery numbers remaining after eliminating applicants seated at schools they've ranked more highly. Applicants' whose most informative disqualification exceeds the cutoff at school s cannot be seated at s because disqualification at a more preferred school implies disqualification at s.

In a match with priorities, the DA propensity score also accounts for the fact that random assignment at s occurs partly as a consequence of not being seated a school preferred to s. Applying these principles in the continuum allows us to describe the DA propensity score as follows:

i) Type Θ_s^n applicants have a DA score of zero because these applicants have worse-than-marginal priority at s.

ii) The probability of assignment at s is $1 - MID_{\theta s}$ for applicants in Θ_s^a because these applicants clear marginal priority at s, but not at higher-ranked choices. Applicants who clear marginal priority at s are guaranteed a seat there if they don't do better. Not doing better means failing to clear $MID_{\theta s}$, the most forgiving cutoff to which they're exposed in the set of schools preferred to s. Since lottery numbers are uniform, this happens occurs with probability $1 - MID_{\theta s}$.

iii) Applicants in Θ_s^c are marginal at s but fail to clear marginal priority at higher-ranked choices. For these applicants to be seated at s they must fail to be seated at a higher-ranked choice and win the competition for seats at s. As for applicants in Θ_s^a, the proportion in Θ_s^c left for consideration at s is $1 - MID_{\theta s}$. Applicants in Θ_s^c are marginal at s, so their status at s is also determined by the lottery cutoff at s. If the cutoff at s, τ_s, falls below the truncation point, $MID_{\theta s}$, no one in this partition finds a seat at s. On the other hand, when τ_s exceeds $MID_{\theta s}$, seats are awarded by drawing from a continuous uniform distribution on $[MID_{\theta s}, 1]$. The resulting assignment probability is therefore $(\tau_s - MID_{\theta s})/(1 - MID_{\theta s})$.

Applying Theorem 1 to the large-market version of Example 2 explains the convergence in type 2 and type 3 propensity scores seen in Figure 1. With no priorities, both types are in Θ_s^c. As we've seen, $MID_{2a} = MID_{3a} = \tau_b = 0.75$, that is, type 2 and 3 students seated at a must have lottery numbers above 0.75. It remains to compute the cutoff, τ_a. Types 2 and 3 compete only with type 4 at a, and are surely beaten out there by type 4s with lottery numbers below 0.75. The remaining 0.25 seats are shared equally between types 2, 3, and 4, going to the best lottery numbers in $[0.75, 1]$, without regard to type. The lottery cutoff at a, τ_a, is therefore

$0.75 + 0.25/3 = 5/6$. Plugging these into equation (3) gives the DA score for types 2 and 3:

$$\begin{aligned}\varphi_a(\theta) &= (1 - MID_{\theta a}) \times \max\left\{0, \frac{\tau_a - MID_{\theta a}}{1 - MID_{\theta a}}\right\} \\ &= (1 - 0.75) \times \max\left\{0, \frac{5/6 - 0.75}{1 - 0.75}\right\} \\ &= \frac{1}{12}.\end{aligned}$$

The score for type 4 is the remaining probability, $1 - (2 \times \frac{1}{12}) = \frac{5}{6}$.

The DA propensity score is a simple function of a small number of intermediate quantities, specifically, $MID_{\theta s}, \tau_s$, and marginal priority status at s and elsewhere. As in Example 2, its common to find that different types have the same marginal priority status and $MID_{\theta s}$, simplifying the score in a manner that facilitates empirical work. In stylized examples, we can easily compute continuum values for these parameters. In real markets with elaborate preferences and priorities, it's natural to use sample analogs for score estimation. As we show below, this generates a consistent estimator of the propensity score for finite markets.

3.3 Estimating the DA Propensity Score

We're interested in the limiting behavior of score estimates based on Theorem 1. The asymptotic sequence for our large-market analysis works as follows: randomly sample n students and their lottery numbers from a continuum economy, described by type distribution F and school capacities, $\{q_s\}$. Call the distribution of types and lottery numbers in this sample F_n. Fix the proportion of seats at school s to be q_s and run DA with these students and schools. Compute $MID_{\theta s}$, τ_s, and partition Θ_s by observing cutoffs $\hat{\mathbf{c}}_n$ and assignments in this single realization, then plug these quantities into equation (3). This generates an estimated propensity score, $\hat{p}_{ns}(\theta)$, constructed by treating a size-n sample economy like its continuum analog. The actual propensity score for this finite economy, computed by repeatedly drawing lottery numbers for the sample of students described by F_n and the set of schools with proportional capacities $\{q_s\}$, is denoted $p_{ns}(\theta)$. We consider the gap between $\hat{p}_{ns}(\theta)$ and $p_{ns}(\theta)$ as n grows. The analysis here makes use of a regularity condition:

Assumption 1. *(First choice support) For any $s \in S$ and priority $\rho \in \{1, ..., K\}$ with $F(\{i \in I : \rho_{is} = \rho\}) > 0$, we have $F(\{i \in I : \rho_{is} = \rho, i \text{ ranks } s \text{ first}\}) > 0$.*

This says that in the continuum economy, every school is ranked first by at least some students in every priority group defined for that school.

In this setup, the propensity score estimated by applying Theorem 1 to data drawn from a single sample and lottery realization converges almost surely to the propensity score generated by repeatedly drawing lottery numbers. This result is presented as a theorem:

Theorem 2. *In the asymptotic sequence described by F_n with proportional school capacities fixed at $\{q_s\}$ and maintaining Assumption 1, the DA propensity score $\hat{p}_{ns}(\theta)$ is a consistent estimator of $p_{ns}(\theta)$ in the following sense: For all $\theta \in \Theta$ and $s \in S$,*

$$|\hat{p}_{ns}(\theta) - p_{ns}(\theta)| \xrightarrow{a.s.} 0.$$

Proof. The proof uses intermediate results, given as lemmas in the theoretical appendix. The first lemma establishes that the vector of cutoffs computed for the sampled economy, $\hat{\mathbf{c}}_n$, converges to the vector of cutoffs in the continuum economy. That is,

$$\hat{\mathbf{c}}_n \xrightarrow{a.s.} \mathbf{c},$$

where $\mathbf{c}$ denotes the continuum economy cutoffs. This result, together with the continuous mapping theorem, implies

$$\hat{p}_{ns}(\theta) \xrightarrow{a.s.} \varphi_s(\theta).$$

In other words, the propensity score estimated by applying Theorem 1 to a sampled finite economy converges to the DA propensity score for the corresponding continuum economy.

A second lemma establishes that for all $\theta \in \Theta$ and $s \in S$,

$$p_{ns}(\theta) \xrightarrow{a.s.} \varphi_s(\theta).$$

since φ_s is a continuous function of cutoffs. That is, the actual (re-randomization-based) propensity score in the sampled finite economy also converges to the propensity score in the continuum economy.[5]

Combining these two results shows that for all $\theta \in \Theta$ and $s \in S$,

$$|\hat{p}_{ns}(\theta) - p_{ns}(\theta)| \xrightarrow{a.s.} |\varphi_s(\theta) - \varphi_s(\theta)| = 0,$$

completing the proof. Since both Θ and S are finite, this also implies uniform convergence, i.e., $\sup_{\theta\in\Theta, s\in S} |\hat{p}_{ns}(\theta) - p_{ns}(\theta)| \xrightarrow{a.s.} 0$. □

Theorem 2 justifies our use of the formula in Theorem 1 to control for student type in empirical work estimating school attendance effects. Specifically, the theorem explains why, as in Example 2, it may be enough to stratify on applicants' most informative disqualification and marginal priority status instead of all possible values of θ when estimating the causal effects of school attendance. Not surprisingly, however, a number of implementation details associated with this strategy remain to be determined. These gaps are filled below.

3.4 Identification

Conditioning on estimates of the propensity score to control for type, we use DPS's first-round charter offers to construct instrumental variables estimates of the effects of charter enrollment on achievement. How should the resulting IV estimates be interpreted? Our IV procedure identifies causal effects for applicants treated when DA produces a charter offer but not otherwise; in the local average treatment effects (LATE) framework of Imbens and Angrist (1994) and Angrist et

[5] See also Azevedo and Leshno (2014), who provide convergence results for the cutoffs and conditional-on-type probabilities of assignment generated by a sequence of stable matchings, showing that the empirical assignment rates for types in a finite market converge to the continuum probability of assignment. The two lemmas in the appendix differ from Azevedo and Leshno (2014)'s results in that they use Assumption 1 and are proved using the extended continuous mapping theorem. The characterization of the DA propensity score in Theorem 1 does not appear to have an analog in the Azevedo and Leshno (2014) framework.

al. (1996), these are charter-offer compliers. IV fails to reveal average causal effects for applicants who decline a first round DA charter offer and are assigned another type of school in round 2 (in the LATE framework, these are never-takers). Likewise, IV methods are not directly informative about the effects of charter enrollment on applicants not offered a charter seat in round 1, but who nevertheless find their way into a charter school in the second round (LATE always-takers).

To flesh out this interpretation and the assumptions on which it rests, let C_i be a charter enrollment indicator and let D_i indicate the offer of a charter seat. These variables indicate attendance and offers at *any* charter school, rather than at a specific school. Since DA produces a single offer, offers of seats at particular schools are mutually exclusive. We can therefore construct D_i by summing individual charter offer dummies. Likewise, the propensity score for this variable, $p_D(\theta) \equiv E[D_i|\theta]$, is obtained by summing the scores for all charter schools to which i has applied.

The population of charter-offer compliers (LATEs) is defined by *potential* treatment status. Potential treatment status (charter enrollment status) is indexed against the DA offer instrument, denoted D_i. In particular, we see potential treatment C_{1i} when D_i is switched on and potential treatment C_{0i} otherwise (both of these are also assumed to exist for all i). Observed treatment is therefore

$$C_i = C_{0i} + (C_{1i} - C_{0i})D_i.$$

Compliers have $C_{1i} - C_{0i} = 1$, an event that happens when $C_{1i} = 1$ and $C_{0i} = 0$.

Causal effects are determined by potential outcomes, indexed against C_i. Initially, we allow for the fact that offers might have a direct effect on outcomes even knowing C_i. This possibility is expressed by writing potential outcomes as $Y_{1i}(d)$ and $Y_{0i}(d)$. This means that when $D_i = d$, we see $Y_{1i}(d)$ if i is treated and we see $Y_{0i}(d)$ otherwise. All four of these potential outcomes are assumed to exist for all i.

Equation (1) implies that conditional on $\theta_i = \theta$, the offer variable, D_i, is independent of potential outcomes and assignments. In a manner analogous to the conditional independence of single-school offers described by equation (1), this can be expressed by writing:

$$\{Y_{1i}(1), Y_{1i}(0), Y_{0i}(1), Y_{0i}(0), C_{1i}, C_{0i}\} \perp\!\!\!\perp D_i|\theta_i; \tag{4}$$

where the vector $\{Y_{1i}(1), Y_{1i}(0), Y_{0i}(1), Y_{0i}(0), C_{1i}, C_{0i}\}$ plays the role of W_i. Likewise, as for single-school offers in equation (2), the propensity score theorem implies

$$\{Y_{1i}(1), Y_{1i}(0), Y_{0i}(1), Y_{0i}(0), C_{1i}, C_{0i}\} \perp\!\!\!\perp D_i|p_D(\theta_i). \tag{5}$$

The conditional independence conditions described by (4) and (5) allow us to estimate causal effects of charter *offers*, the treatment indicated by D_i. In practice, however, we're interested in the effects of charter *attendance*, the treatment indicated by C_i.

Identification of average causal effects of charter school attendance requires an exclusion restriction. Specifically, we assume

$$Y_{ji}(1) = Y_{ji}(0) \equiv Y_{ji}; j = 0, 1.$$

In other words, charter offers are assumed to be unrelated to outcomes for applicant i once we know whether this applicant attended a charter school. The exclusion restriction allows us to

replace the four double-index potential outcomes in (4) and (5) with two single-index potential outcomes, Y_{1i} and Y_{0i}.

The case for the exclusion restriction is less immediate than that for conditional independence. We might worry, for example, that outcomes are affected by lottery numbers even for applicants whose charter status is unchanged by the lottery (that is, for always-takers and never-takers). Denver's second round allocates any remaining school seats in an *ad hoc* school-by-school application process, unrelated to lottery numbers drawn in the first round. But lottery numbers can nevertheless affect the second round indirectly by changing opportunities. Consider, for example, a skittish charter applicant who chooses a popular non-charter option when $D_i = 0$ in round 1. Fearing the long charter school day and having applied to charter schools only to satisfy his mother, this applicant also goes non-charter if his $D_i = 1$. But in this case, having been offered a charter seat in round 1, he must settle for a less desirable and perhaps lower-quality non-charter option in round 2. This violates the exclusion restriction if $Y_{0i}(1) \neq Y_{0i}(0)$. We must therefore either assume away within-sector differences in potential outcomes, or introduce a finer-grained parameterization of school sector effects. The latter approach is explored in Section 4.6, below.

In addition to the conditional independence and exclusion restrictions, we also assume that, conditional on the propensity score, charter offers cause charter enrollment for at least some students, and that charter offers can only make charter enrollment more likely, so that $C_{1i} \geq C_{0i}$ for all i. Given these assumptions, the conditional-on-score IV estimand is a conditional average causal affect for compliers, that is:

$$\frac{E[Y_i|D_i = 1, p_D(\theta_i) = x] - E[Y_i|D_i = 0, p_D(\theta_i) = x]}{E[C_i|D_i = 1, p_D(\theta_i) = x] - E[C_i|D_i = 0, p_D(\theta_i) = x]} = E[Y_{1i} - Y_{0i}|p_D(\theta_i) = x, C_{1i} > C_{0i}], \quad (6)$$

where $p_D(\theta_i)$ is the charter-offer propensity score associated with applicant i's type and x indexes values in the support of $p_D(\theta)$.

3.5 Estimation

In view of the fact that (6) generates a distinct causal effect for each score value, it's natural to consider parsimonious models that use data from all propensity-score cells to estimate a single average causal effect. We accomplish this by estimating a 2SLS specification with first and second stage equations that can be written

$$C_i = \sum_x \gamma(x) d_i(x) + \delta D_i + X_i'\lambda + \nu_i, \quad (7)$$

$$Y_i = \sum_x \alpha(x) d_i(x) + \beta C_i + X_i'\mu + \epsilon_i, \quad (8)$$

where the $d_i(x)$'s are dummies indicating values of $p_D(\theta_i)$, indexed by x, and $\gamma(x)$ and $\alpha(x)$ are the associated "score effects" in the first and second stages. The coefficient δ in (7) is the first-stage effect of charter offers on charter enrollment, while the coefficient β in (8) is the causal effect of interest. These first and second stage equations include baseline covariates, X_i, to increase precision and adjust for any chance imbalances in student characteristics.

As a check on the 2SLS specification, we also report semiparametric estimates of $E[Y_{1i} - Y_{0i}|C_{1i} > C_{0i}]$. In contrast with the additive 2SLS setup, the semiparametric procedure requires

only correct specification of the propensity score to generate a single average causal effect for all compliers. The semiparametric strategy is founded on Abadie (2003)'s observation that the conditional independence and exclusion restrictions imply:

$$E[Y_{0i}|C_{1i} > C_{0i}] = \frac{1}{Pr(C_{1i} > C_{0i})} E\left[\frac{C_i Y_i (D_i - p_D(\theta_i))}{(1 - p_D(\theta_i))p_D(\theta_i)}\right],$$
$$E[Y_{1i}|C_{1i} > C_{0i}] = \frac{1}{Pr(C_{1i} > C_{0i})} E\left[\frac{(1 - C_i) Y_i ((1 - D_i) - (1 - p_D(\theta_i)))}{(1 - p_D(\theta_i))p_D(\theta_i)}\right].$$

Subtracting and rearranging, we have:

$$E[Y_{1i} - Y_{0i}|C_{1i} > C_{0i}] = \frac{1}{Pr(C_{1i} > C_{0i})} E\left[\frac{Y_i (D_i - p_D(\theta))}{(1 - p_D(\theta_i))p_D(\theta_i)}\right]. \quad (9)$$

The first stage in this case, $P[C_{1i} > C_{0i}]$, is constructed using

$$P[C_{1i} > C_{0i}] = E\left[\frac{C_i (D_i - p_D(\theta_i))}{(1 - p_D(\theta_i))p_D(\theta_i)}\right]. \quad (10)$$

The semi-parametric IV estimator used here is the sample analog of the right hand side of (9) divided by the sample analog of (10).[6]

4 School Effectiveness in Denver

Since the 2011 school year, DPS has used DA to assign students to most schools in the district, a process known as SchoolChoice. Denver school assignment involves two rounds, but only the first round uses DA. Our analysis therefore focuses on the initial round.

In the first round of SchoolChoice, parents can rank up to five schools of any type, including traditional public schools, magnet schools, innovation schools, and most charters. A neighborhood school is also ranked automatically (if a student has a neighborhood school, the district adds his neighborhood school to his choice list as the last choice). Schools ration seats using a mix of priorities and a single lottery number. Priorities vary across schools and typically involve siblings and neighborhoods. Seats may be reserved for a certain number of subsidized-lunch students and for children of school staff. Reserved seats are allocated by splitting schools and assigning the highest priority status to students in the reserved group at one of the sub-schools created by a split. Match participants can only qualify for seats in a single grade.

The DPS match distinguishes between groups of seats at a given school, known as "buckets." Buckets in the same school have distinct priorities and capacities. DPS converts applicants' preferences over schools into preferences over buckets, splitting off separate sub-schools for each. The upshot for our purposes is that DPS's version of DA assigns seats at the sub-schools determined by seat reservation policies and buckets rather than schools, while the relevant propensity score captures the probability of offers at sub-schools. The discussion that follows refers to propensity scores for schools, with the understanding that the fundamental unit of assignment is a bucket, from which assignment rates to schools have been constructed.[7]

[6] Covariates are incorporated in the semiparametric estimation procedure by adding X_i to a logit model for $E[D_i|X_i, p_D(\theta_i)]$ and using fitted values from this instead of estimates of $p_D(\theta_i)$ in equations (9) and (10).

[7] DPS modifies traditional DA mechanism by recoding the lottery numbers of all siblings applying to the same

4.1 Computing the DA Propensity Score

The score estimates used as controls in equations (7) and (8) were constructed three ways. The first is a simulation-based benchmark: we ran DA for one million lottery draws and recorded the proportion of draws in which applicants of a given type in our fixed DPS sample were seated at each school.[8] By a conventional law of large numbers, this simulated score converges to the actual finite-market score as the number of draws increases. In practice, of course, the number of replications is far smaller than the number of possible lottery draws, so the simulated score takes on more values than we'd expect to see for the actual score. For applicants with a simulated score strictly between zero and one, the simulated score takes on more than 1,100 distinct values (with fewer than 1,300 types in this sample). Because many simulated score values are exceedingly close to one another (or to 0 or 1) some of the estimators that control for the simulated score use values that have been rounded.

We're particularly interested in taking advantage of the DA score defined in Theorem 1. This theoretical result is used for propensity score estimation in two ways. The first, which we label a "formula" calculation, applies equation (3) directly to the DPS data. Specifically, for each applicant type, school, and entry grade, we identified marginal priorities, and applicants were allocated by priority status to either Θ_s^n, Θ_s^a, or Θ_s^c. The DA score, $\varphi_s(\theta)$ is then estimated by computing the sample analog of $MID_{\theta s}$ and τ_s in the DPS assignment data and plugging these into equation (3).

The bulk of our empirical work uses a second application of Theorem 1, which also starts with marginal priorities, MIDs, and cutoffs in the DPS data. This score estimate, however, is given by the empirical offer rate in cells defined by these variables. This score estimate, which we refer to as a "frequency" calculation, is closer to an estimated score of the sort discussed by Abadie and Imbens (2012) than is the formula score, which ignores realized assignment rates. The large-sample distribution theory in Abadie and Imbens (2012) suggests that conditioning on an estimated score based on realized assignment rates may increase the efficiency of score-based estimates of average treatment effects.

Propensity scores for school offers tell us the number of applicants subject to random assignment at each DPS charter school.[9] These counts, reported in columns 3-5 of Table 1 for the three

school to be the best random number held by any of them. This modification (known as "family link") changes the allocation of only about 0.6% of students from that generated by standard DA. Our analysis incorporates family link by defining distinct types for linked students.

[8] Calsamiglia et al. (2014) and Agarwal and Somaini (2015) simulate the Boston mechanism as part of an effort to estimate preferences in a structural model of latent preferences over schools.

[9] The data analyzed here come from files containing the information used for first-round assignment of students applying in the 2011-12 school year for seats the following year (this information includes preference lists, priorities, random numbers, assignment status, and school capacities). School-level scores were constructed by summing scores for all component sub-schools used to implement seat reservation policies and to define buckets. Our empirical work also uses files with information on October enrollment and standardized scores from the Colorado School Assessment Program (CSAP) and the Transitional Colorado Assessment Program (TCAP) tests, given annually in grades 3-10. A data appendix describes these files and the extract we've created from them. For our purposes, "Charter schools" are schools identified as "charter" in DPS *2012-2013 SchoolChoice Enrollment Guide* brochures and not identified as "intensive pathways" schools, which serve students who are much older than typical for their grade.

different score estimators, range from none to over 300. The proportion of applicants subject to random assignment varies markedly from school to school. This can be seen by comparing the count of applicants subject to random assignment with the total applicant count in column 1. The randomized applicant count calculated using frequency and formula score estimates are close, but some differences emerge when a simulated score is used.[10]

Column 5 of Table 1 also establishes the fact that at least some applicants were subject to random assignment at every charter except for the Denver Language School, which offered no seats. In other words, every school besides the Denver Language School had applicants with a simulated propensity score strictly in the unit interval. Three schools for which the simulated score shows very few randomized applicants (Pioneer, SOAR Oakland, Wyatt) have an empirical offer rate of zero, so the frequency version of the DA propensity score is zero for these schools (applicant counts based on intervals determined by DA frequency and formula scores appear in columns 3 and 4).

DA produces random assignment of seats for students ranking charters *first* for a much smaller set of schools. This can be seen in the last column of Table 1, which reports the number of applicants with a simulated score strictly between zero and one, who also ranked each school first. The reduced scope of first-choice randomization is important for our comparison of strategies using the DA propensity score with previously-employed IV strategies using first-choice instruments. First-choice instruments applied to the DPS charter sector necessarily ignore many schools. Note also that while some schools had only a handful of applicants subject to random assignment, over 1,400 students were randomized in the charter sector as a whole.

The number of applicants randomized at particular schools can be understood further using Theorem 1. Why did STRIVE Prep - GVR have 116 applicants randomized, even though Table 1 shows that no applicant with non-degenerate offer risk ranked this school first? Random assignment at GVR is a consequence of the many GVR applicants randomized by admissions offers at schools they'd ranked more highly. This and related determinants of offer risk are detailed in Table 2, which explores the anatomy of the DA propensity score for 6th grade applicants to four middle schools in the STRIVE network. In particular, we see (in column 8 of the table) that all randomized GVR applicants were randomized by virtue of having $MID_{\theta s}$ inside the unit interval, with no one randomized at GVR's own cutoff (column 7 counts applicants randomized at each school's cutoff).

In contrast with STRIVE's GVR school, few applicants were randomized at STRIVE's Highland, Lake, and Montbello campuses. This is a consequence of the fact that most Highland, Lake, and Montbello applicants were likely to clear marginal priority at these schools (having $\rho_{\theta s} < \rho_s$), while having values of $MID_{\theta s}$ mostly equal to zero or 1, eliminating random assignment at schools ranked more highly. Interestingly, the Federal and Westwood campuses are the only STRIVE schools to see applicants randomized around the cutoff in the school's own marginal priority group. We could therefore learn more about the impact of attendance at Federal and Westwood by changing the cutoff there (e.g., by changing capacity), whereas such a

[10]The gap here is probably due to our treatment of family link. The Blair charter school, where the simulated score randomization count is farthest from the corresponding DA score counts, has more applicants with family link than any other school. Unlike our DA score calculation, which ignores family link, the simulated score accommodates family link by assigning a unique type to every student affected by a link.

change would be of little consequence for evaluations of the other schools.

Table 2 also documents the weak connection between applicant randomization counts and a naive definition of over-subscription based on school capacity. In particular, columns 2 and 3 reveal that four out of six schools described in the table ultimately made fewer offers than they had seats available (far fewer in the case of Montbello). Even so, assignment at these schools was far from certain: they contribute to our score-conditioned charter school impact analysis.

A broad summary of DPS random assignment appears in Figure 2. Panel (a) captures the information in columns 3 and 6 of Table 1 by plotting the number of first-choice applicants subject to randomization as black dots, with the total randomized at each school plotted as an arrow pointing up from these dots (schools are indexed on the x-axis by their capacities). This representation highlights the dramatic gains in the number of schools and the precision with which they can be studied as a payoff to our full-information approach to the DA research design. These benefits are not limited to the charter sector, a fact documented in Panel (b) of the figure, which plots the same comparisons for non-charter schools in the DPS match.

4.2 DPS Data and Descriptive Statistics

The DPS population enrolled in grades 3-9 in the Fall of 2011 is roughly 60% Hispanic, a fact reported in Table 3, along with other descriptive statistics. We focus on grades 3-9 in 2011 because outcome scores come from TCAP tests taken in grades 4-10 in the spring of the 2012-13 school year.[11] The high proportion Hispanic makes DPS an especially interesting and unusual urban district. Not surprisingly in view of this, almost 30 percent of DPS students have limited English proficiency. Consistent with the high poverty rates seen in many urban districts, three quarters of DPS students are poor enough to qualify for a subsidized lunch. Roughly 20% of the DPS students in our data are identified as gifted, a designation that qualifies them for differentiated instruction and other programs.

Nearly 11,000 of the roughly 40,000 students enrolled in grades 3-9 in Fall 2011 sought to change their school for the following year by participating in the assignment, which occurs in the spring. The sample participating in the assignment, described in column 2 of Table 3, contains fewer charter school students than appear in the total DPS population, but is otherwise demographically similar. It's also worth noting that our impact analysis is limited to students enrolled in DPS in the baseline (pre-assignment) year of 2011. The sample described in column 2 is therefore a subset of that described in column 1. The 2012 school assignment, which also determines the propensity score, includes the column 2 sample plus new entrants.

Column 3 of Table 3 shows that of the nearly 11,000 DPS-at-baseline students included in the assignment, almost 5,000 ranked at least one charter school. We refer to these students as charter applicants; the estimated charter attendance effects that follow are for subsets of this applicant group. DPS charter applicants have baseline achievement levels and demographic characteristics broadly similar to those seen district-wide. The most noteworthy feature of the charter applicant sample is a reduced proportion white, from about 19% in the centralized assignment to a little over 12% among charter applicants. It's also worth noting that charter applicants have baseline

[11] Grade 3 is omitted from the outcome sample because 3rd graders have no baseline test.

test scores close to the DPS average. This contrasts with the modest positive selection of charter applicants seen in Boston (reported in Abdulkadiroğlu et al. 2011).

A little over 1,400 charter applicants have a frequency estimate of the probability of charter assignment between zero and one; the count of applicants subject to random assignment rises to about 1,500 when the score is estimated by simulation. Charter applicants subject to random assignment are described in columns 4 and 6 of Table 3. Although only about 30% of charter applicants were randomly assigned a charter seat, these students look much like the full charter applicant pool. The main difference is a higher proportion of applicants of randomized applicants originating at a charter school (that is, already enrolled at a charter at the time they applied for seats elsewhere). Columns 5 and 7, which report statistics for the subset of the randomized group that enrolls in a charter school, show slightly higher baseline scores among charter students.

4.3 Score-Based Balance

Conditional on the propensity score, applicants offered a charter seat should look much like those not offered a seat. Moreover, because offers are randomly assigned conditional on the score, we expect to see conditional balance in all applicant characteristics and not just for the variables that define an applicant's type. We assess the balancing properties of the DA propensity score using simulated expectations. Specifically, drawing lottery numbers 400 times, we ran DA and computed the DA propensity score each time, and then computed average covariate differences by offer status. The balance analysis begins with uncontrolled differences in average applicant characteristics, followed by regression-adjusted differences that put applicant characteristics on the left-hand side of regression models like equation (7), omitting the covariate controls, X_i.

Uncontrolled comparisons by offer status, reported in columns 1 and 2 of Table 4, show large differences in average student characteristics, especially for variables related to preferences. For instance, across 400 lottery draws, those not offered a charter seat ranked an average of 1.4 charters, but this figure increases by almost half a school for applicants who were offered a charter seat. Likewise, while fewer than 30% of those not offered a charter seat had ranked a charter school first, the probability applicants ranked a charter first increases to over 0.9 (that is, 0.29+0.62) for those offered a charter seat. Column 2 also reveals important demographic differences by offer status; Hispanic applicants, for example, are substantially over-represented among those offered a charter seat.[12]

Conditioning on frequency estimates of the DA propensity score reduces differences by offer status markedly. This can be seen in columns 3-5 of Table 4. The first set of conditional results, which come from regression models with linear control for the propensity score rather than dummies, show virtually no difference by offer status in the odds a charter is ranked first or that an applicant is Hispanic. Offer gaps in other application and demographic variables are also much reduced in this specification. Columns 4 and 5 of the table show that non-parametric control for the DA propensity score (implemented by dummying all score values in the unit interval, with an average of 39 values across simulations when rounded to nearest hundredth and an average

[12]Table 4 omits standard errors because the only source of uncertainty here is the modest simulation error arising from the fact that we've drawn lottery numbers 400 instead of infinitely many times.

47 without rounding) reduces offer gaps even further. These results establish that a single DPS applicant cohort is large enough for the DA propensity score to eliminate selection bias.[13]

Columns 6-8 of Table 4, which report estimated offer gaps conditional on a simulated propensity score, show that the simulated score does a better job of balancing treatment and control groups than does the DA score. Differences by offer status conditional on the simulated score, whether estimated linearly or with nonparametric controls, appear mostly in the third decimal place. This reflects the fact that simulation recovers the actual finite-market propensity score (up to simulation error), while the DA propensity score is an asymptotic approximation that should be expected to provide perfect treatment-control balance only in the limit. It's worth noting, however, that the simulated score starts with 1,148 unique values. As a practical matter, the simulated score must be smoothed to accommodate non-parametric control. Rounding to the nearest hundredth leaves us with 51 points of support, close to the number of support points seen for the DA score. Rounding to the nearest ten-thousandth leaves 121 points of support. Finer rounding produces noticeably better balance for the number-of-schools-ranked variable.

Our exploration of score-based balance is rounded out with the results from a traditional balance analysis such as would be seen in analyses of a randomized trial. Specifically, Table 5 documents balance for the DPS match by reporting the usual t and F-statistics for offer gaps in covariate means. Again, we look at balance conditional on propensity scores for applicants with scores strictly between 0 and 1. As can be seen in Table 5a, application covariates are well-balanced by non-parametric control for either DA or simulated score estimates (linear control for the DA propensity score leaves a significant gap in the number of charter schools ranked).[14]

Table 5a also demonstrates that full control for type leaves us with a much smaller sample than does control for the propensity score: models with full type control are run on a sample of size 301, a sample size reported in the last column of the table. Likewise, the fact that saturated control for the simulated score requires some smoothing can be see in the second last column showing the reduced sample available for estimation of models that control fully for a simulated score rounded to the nearest ten-thousandth.

Not surprisingly, a few significant imbalances emerge in balance tests for the longer list of baseline covariates, reported in Table 5b. Here, the simulated score seems to balance characteristics somewhat more completely than does the DA score, but the F-statistics (reported at the bottom of the table) that jointly test balance of all baseline covariates fail to reject the null hypothesis of conditional balance for any specification reported.

Baseline score gaps as large as -0.1σ appear in some of the comparisons at the bottom of the table. The fact that these gaps are not mirrored in the comparisons in Table 4 suggests the differences in Table 5 are due to chance. Still, we can mitigate the effect of chance differences on 2SLS estimates of charter effects by adding baseline score controls (and other covariates) to our empirical models. The inclusion of these additional controls also has the salutary effect of making

[13] Appendix Table B4 provides a computational proof of Theorem 2 by reporting offer gaps of the sort shown in Table 4 for scaled-up versions of the DPS economy. Doubling the number of applicants and seats at each school in the DPS market pushes conditional gaps down markedly, and multiples of 4 and 8 make these small gaps even smaller.

[14] Table 5 reports the results controlling for frequency estimates of the DA propensity score and the simulated propensity score. Balance results using formula estimates of the score appear in Appendix Table B3.

the 2SLS estimates of interest considerably more precise (covariates include dummies for grade tested, gender, origin school charter status, race, gifted status, bilingual status, subsidized lunch eligibility, special education, limited English proficient status, and baseline test scores; baseline score controls are responsible for most of the resulting precision gain). Finally, it's worth noting that the imbalance left after conditioning on the DA propensity score turns out to matter little for the 2SLS estimates we're ultimately after.

Modes of Inference

Econometric inference typically tries to quantify the uncertainty due to *random sampling*. What then, to make of the fact that the analysis reported here uses data on all DPS applicants from 2012? On one hand, we might imagine that the applicants we happen to be studying constitute a random sample from some larger population of possible applicants. At the same time, the statistical uncertainty in our empirical work can also be seen as a consequence of *random assignment*: we see only a single lottery draw for each applicant, one of many possibilities even when the sample of applicants is viewed as fixed.

In an effort to determine whether the distinction between sampling inference and randomization inference matters for our purposes, we computed randomization p-values by repeatedly drawing lottery numbers and calculating offer gaps in covariates conditional on the simulated propensity score. Regression conditioning on the simulated score produces near-perfect balance in Table 4 so this distribution is what we should expect to see under the null hypothesis of no difference by treatment assignment. Randomization p-values are therefore given by quantiles of the t-statistics in the distribution resulting from these repeated draws.

The p-values associated with conventional robust t-statistics for covariate balance turn out to be close to the corresponding randomization p-values. For the number of charter schools an applicant has ranked, for example, the conventional p-value for balance is 0.885 while the corresponding randomization p-value is 0.850. This is consistent with a classic result on the asymptotic equivalence of randomization and sampling tests for differences in means (see, e.g., 15.2 in Lehmann and Romano 2005).

Abadie et al. (2014) generalize results on the large-sample equivalence of randomization and sampling inference to cover regression estimates of treatment effects and tests for covariate balance of the sort reported here. If the regression function is linear and the regression of treatment on controls is linear, the usual robust covariance matrix associated with random sampling is asymptotically valid for the sampling distribution induced by random assignment. The treatment in our case is an offer dummy, while the controls are dummies or a linear model for the propensity score. The second of these requirements holds here when the controls fully saturate the propensity score (ignoring any additional covariates). The first requires constant offer effects given a saturated model for the score. The models estimated here don't quite satisfy these conditions (they're not fully saturated) but do not seem to be so far off that this matters for inference.

A related issue arises from the fact that the empirical strategy used here conditions on estimates of the propensity score (the simulated score is also an estimate since it's based on a finite number of draws). As noted by Hirano et al. (2003) and Abadie and Imbens (2012), conditioning

on an estimated as opposed to a non-stochastic known score may affect sampling distributions of the resulting estimated causal effects. We therefore checked conventional large-sample p-values against randomization p-values for the reduced-form charter offer effects associated with the 2SLS estimates discussed in the next section. Robust asymptotic sampling formulas again generate p-values close to a randomization-inference benchmark, regardless of how the score behind these estimates was constructed. In view of these findings, we rely on the usual robust standard errors and test statistics for inference about 2SLS estimates of treatment effects.[15]

4.4 Effects of Charter Enrollment

2SLS estimates of charter attendance effects are remarkably similar to the corresponding semiparametric estimates. This is apparent in Table 6, which compares 2SLS estimates of models with additive score controls to semiparametric estimates of average treatment effects constructed using three versions of the score. Compare, for example, frequency-score-controlled 2SLS estimates of effects on math and reading of 0.496 and 0.127 with semiparametric estimates around 0.44 and 0.11. At the same time, standard errors for the semiparametric estimates are higher than those for 2SLS (semiparametric precision is estimated using a Bayesian bootstrap that randomly reweights observations; see Shao and Tu (1995) for an introduction). Semiparametric estimates weighted using a simulated score are especially imprecise. The similarity of 2SLS and semiparametric estimates and the relative simplicity of 2SLS estimation leads us to report only 2SLS estimates in what follows.[16]

A DA-generated charter offer boosts charter school attendance rates by about 0.4. These first stage estimates, shown in the first row of Table 7, are computed by estimating equation (7). The first stage of 0.4 reflects the fact that many charter applicants who are not offered a seat in the SchoolChoice first round ultimately find their way into a charter school by applying to schools directly in the second round (specifically, 43% of the charter applicants analyzed in Table 7 are always-takers who enroll in charters even without a first-round charter offer, while fewer than 20% of the analysis sample are never-takers who decline charter offers). First-stage estimates of around 0.68 computed without score controls, shown in column 4 of the table, are clearly biased upwards.[17]

[15]Online Appendix Table B2 reports conditional-on-score estimates of attrition differentials by offer status. Here, we see marginally significant gaps on the order of 4-5 points when estimated conditional on the DA propensity score. Attrition differentials fall to a statistically insignificant 3 points when estimated conditional on a simulated score. The estimated charter attendance effects discussed below are similar when computed using either type of score control, so it seems unlikely that differential attrition is a source of bias in our 2SLS estimates.

[16]2SLS also obviates the need for judgements regarding bootstrap methods or implementation. We found, for example, that a conventional nonparametric bootstrap for the semiparametric estimators requires trimming or tuning to eliminate the influence of occasional small first stage estimates.

[17]The estimates reported in this table control for baseline test scores and the covariates described earlier. These extra controls are not necessary for consistent causal inference but their inclusion increases precision (Estimates without covariates appear in the appendix). Estimates here are for scores in grades 4-10. The pattern of results in an analysis that separates high schools from middle and elementary schools is similar. The sample used for IV estimation is limited to charter applicants with the relevant propensity score in the unit interval, for which score cells have offer variation in the data at hand (these restrictions amount to the same thing for the frequency score). The OLS estimation sample includes charter applicants, ignoring score- and cell-variation restrictions.

2SLS estimates of charter attendance effects on test scores, reported below the first-stage estimates in Table 7, show remarkably large gains in math, with smaller effects on reading. The math gains reported here are similar to those found for charter students in Boston (see, for example, Abdulkadiroğlu et al. 2011). Previous lottery-based studies of charter schools likewise report substantially larger gains in math than in reading. Here, we also see large and statistically significant gains in writing scores.

Importantly for our methodological agenda, the estimated charter attendance effects reported in Table 7 are largely invariant to whether the propensity score is estimated by simulation or by a frequency or formula calculation that uses Theorem 1. Compare, for example, math impact estimates of 0.496, 0.524, and 0.543 using frequency-, formula-, and simulation-based score controls, all estimated with similar precision. This alignment validates the use of Theorem 1 to control for applicant type.

Estimates that omit propensity score controls highlight the risk of selection bias in a naive 2SLS empirical strategy. This is documented in column 4 of Table 7, which shows that 2SLS estimates of math and writing effects constructed using DA offer instruments while omitting propensity score controls are too small by about half. A corresponding set of OLS estimates without propensity score controls, reported in column 5 of the table, also tends to underestimate the gains from charter attendance attendance. The results in column 6 show that adding score controls to the OLS model pulls the estimates up a little, but a substantial gap between between these and the corresponding set of 2SLS estimates remains.

4.5 Alternative IV Strategies

We're interested in comparing 2SLS estimates constructed using a DA offer dummy as an instrument while controlling for the DA propensity score with suitably-controlled estimates constructed using first-choice and qualification instruments. As noted in Section 3.4, we expect DA-offer instruments to yield a precision gain and to increase the number of schools represented in the estimation sample relative to these two previously-employed IV strategies.[18]

Let $R(\theta_i)$ be a variable that uniquely identifies the charter school that applicant i ranks first, along with his priority status at this school, defined for applicants whose first choice is indeed a charter school. $R(\theta_i)$ ignores other schools that might have been ranked. The first-choice strategy is implemented by the following 2SLS setup:

$$Y_i = \sum_x \alpha(x) d_i(x) + \beta C_i + \epsilon_i,$$
$$C_i = \sum_x \gamma(x) d_i(x) + \delta D_i^f + \nu_i,$$

where the $d_i(x)$'s are dummies indicating values of $R(\theta_i)$, indexed by x, and $\gamma(x)$ and $\alpha(x)$ are the associated "risk set effects" in the first and second stages. The first-choice instrument, D_i^f,

[18]Studies using first-choice instruments to evaluate schools in districts with centralized assignment include Abdulkadiroğlu et al. (2013), Deming (2011), Deming et al. (2014), and Hastings et al. (2009). First-choice instruments have also been used with decentralized assignment mechanisms (Abdulkadiroğlu et al. (2011), Cullen et al. (2006), Dobbie and Fryer (2011), and Hoxby et al. (2009)). Dobbie and Fryer (2014), Lucas and Mbiti (2014), and Pop-Eleches and Urquiola (2013) use qualification instruments.

is a dummy variable indicating i's qualification at his or her first-choice school. In other words,

$$D_i^f = 1[\pi_{is} \leq c_s \text{ for charter } s \text{ that } i \text{ has ranked first}].$$

First choice qualification is the same as first choice offer since under DA, applicants who rank a first are offered a seat there if and only if they qualify at a.

The qualification strategy expands the sample to include all charter applicants, with the risk sets ($R(\theta_i)$) for qualification instruments identifying the set of all charter schools that i ranks, along with his or her priority status at each of these schools. In this case, $R(\theta_i)$ ignores the order in which schools are ranked, coding only their identities, but priorities are associated with schools.[19] The qualification instrument, D_i^q, indicates qualification at *any* charter he or she has ranked. In other words,

$$D_i^q = 1[\pi_{is} \leq c_s \text{ for at least one charter } s \text{ that } i \text{ has ranked}].$$

In large markets, the instruments D_i^f and D_i^q are independent of type conditional on $R(\theta_i)$; see Appendix A.5 for details.

A primary source of inefficiency in the first-choice and qualification strategies is apparent in Panel A of Table 8. This panel reports two sorts of first stage estimates for each instrument: the first of these regresses a dummy indicating any charter *offer*–that is, our DA charter offer instrument, D_i–on each of the three instruments under consideration. A regression of D_i on itself necessarily produces a coefficient of one. By contrast, a first-choice offer boosts the probability of any charter offer by only around 0.77 in the sample of those who have ranked a charter first. This reflects the fact that, while anyone receiving a first choice charter offer has surely been offered a charter seat, roughly 23% of the sample ranking a charter first is offered a charter seat at schools other than their first choice. The relationship between D_i^q and charter offers is even weaker, at around 0.48. This reflects the fact that for schools below the one ranked first, charter qualification is not sufficient for a charter offer.

The diminished impact of the two alternative instruments on charter offers translates into a weakened first stage for charter *enrollment*. The best case scenario, using all DA-generated offers (that is, D_i) as a source of quasi-experimental variation, produces a first stage of around 0.41. But first-choice offers boost charter enrollment by just 0.32, while qualification anywhere yields a charter enrollment gain of only 0.18. As always with comparisons of IV strategies, the size of the first stage is a primary determinant of relative precision.

At 0.071, the standard error of the DA-offer estimate is markedly lower than the standard error of 0.102 yielded by a first-choice strategy and well below the standard error of 0.149 generated by qualification instruments. In fact, the precision loss here is virtually the same as the decline in the intermediate first stages recorded in the first row of the table (compare 0.774 with $0.071/0.102 = 0.696$ and 0.476 with $0.071/0.149 = 0.477$). The loss here is substantial: columns 4 and 5 show the sample size increase needed to undo the damage done by a smaller first stage for each alternative instrument.[20]

[19] For example, an applicant who ranks A and B with marginal priority only at A is distinguished from an applicant who ranks A and B with marginal priority only at B.

[20] The sample used to construct the estimates in columns 1-3 of Table 8 is limited to those who have variation in the instrument at hand conditional on the relevant risk sets controls.

Only half as many schools are represented in the first-choice analysis sample as in the DA sample (At 24, the number of schools in the qualification sample is closer to the full complement of 30 schools available for study with DA offers). First-choice analyses lose schools because many lotteries fail to randomize first-choice applicants (as seen in Table 1). It's therefore interesting to note that the first-choice estimate of effects on math and reading scores are noticeably larger than the estimates generated using DA offer and qualification instruments (compare the estimate of 0.5 using DA offers with estimates of 0.6 and 0.41 using first-choice and qualification instruments). This finding may reflect an advantage for those awarded a seat at their first choice school (Hastings et al. 2009; Deming 2011; Deming et al. 2014 find a general "first choice advantage" in analyses of school attendance effects.) By contrast, the DA offer instrument yields an estimand that is more representative of the full complement of charter schools in the match. In the same spirit, it's worth noting that the first-choice and qualification IV samples include no 10th graders.

4.6 Charter School Effects with a Mixed Counterfactual

The 2SLS estimates in Tables 7 and 8 contrast charter outcomes with potential outcomes generated by attendance at a mix of traditional public schools and schools from other non-charter sectors. We'd like to simplify this mix so as to produce something closer to a pure sector-to-sector comparison. Allowance for more than one treatment channel also addresses concerns about charter-offer-induced changes in counterfactual outcomes that might cause violations of the exclusion restriction.

Our first step in this effort is to describe the distribution of non-charter school choices for applicants who were and weren't offered a charter seat in the DPS assignment. We then identify the distribution of counterfactual (non-charter) school sectors for the group of charter-lottery compliers. Finally, we use the DA mechanism to jointly estimate causal effects of attendance at schools in different sectors, thereby making the non-charter counterfactual in our 2SLS estimates more homogeneous.

The analysis here builds on a categorical variable, W_i, capturing school sector in which i is enrolled. Important DPS sectors besides the charter sector are traditional public schools, innovation schools, magnet schools, and alternative schools. Innovation and magnet schools are managed by DPS. Innovation schools design and implement innovative practices meant to improve student outcomes (for details and a descriptive evaluation of innovation schools, see Connors et al. 2013). Magnet schools serve students with particular styles of learning. Alternative schools serve mainly older students struggling with factors that may prevent them from succeeding in a traditional school environment. Smaller school sectors include a single charter middle school outside the centralized DPS assignment process (now closed) and a private school contracted to serve DPS students.

The distribution of enrollment sectors for students who do and don't receive a charter offer are described in the first two columns of Table 9. These columns show a charter enrollment rate of 87% in the group offered a charter seat, along with substantial but much smaller charter enrollment in the non-offered group.[21] Perhaps surprisingly, only around 41% of those not

[21] As noted in the discussion of the first stage estimates in Table 7, applicants unhappy with the offer they've receive in the first round of SchoolChoice can apply to schools individually in a second round.

offered a charter seat enroll in a traditional public schools, with the rest of the non-offered group distributed over a variety of school types. Innovation schools are the leading non-charter alternative to traditional public schools. Innovation schools operate under an innovation plan that waives some provisions of the relevant collective bargaining agreements (for a descriptive evaluation of these schools, see Connors et al. 2013).[22]

The sector distribution for non-offered applicants with non-trivial charter risk appears in column 3 of Table 9, alongside the sum of the non-offered mean and a charter-offer treatment effect on enrollment in each sector in column 4. These extended first-stage estimates, computed by putting indicators $1(W_i = j)$ on the left-hand side of equation (7), control for the DA propensity score and therefore have a causal interpretation. The number of applicants not offered a seat who end up in a charter school is higher for those with non-trivial charter offer risk than in the full applicant sample, as can be seen by comparing columns 3 and 1. The charter enrollment first stage that's implicit in the column 4-vs-3 comparison matches the first stage in Table 7. First stages for other sectors show charter offers sharply reduce innovation school enrollment as well as reducing enrollment in traditional public schools.

The 2SLS estimates reported in Table 7 capture causal effect for charter lottery compliers. We describe the distribution of school sectors for compliers by defining *potential* sector enrollment variables, W_{1i} and W_{0i}, indexed against charter offers, D_i. Potential and observed sector variables are related by

$$W_i = W_{0i} + (W_{1i} - W_{0i})D_i.$$

In the population of charter-offer compliers, $W_{1i} = \textit{charter}$ for all i: by definition, charter-offer compliers attend a charter school when the DPS assignment offers them the opportunity to do so. Here, we're interested in $E[1(W_{0i} = j)|C_{1i} > C_{0i}]$, that is, the sector distribution for charter-offer compliers in the scenario where they aren't offered a charter seat. We refer to this distribution as describing *enrollment destinies* for compliers.

Enrollment destinies are marginal potential outcome distributions for compliers. As shown by Abadie (2002), these are identified by a simple 2SLS estimand. The details of our implementation of this identification strategy follow those in Angrist et al. (2015), with the modification that instead of estimating marginal potential outcome densities for a continuous variable, the outcomes of interest here are Bernoulli.[23]

Column 5 of Table 9 reveals that only about half of charter lottery compliers are destined to end up in a traditional public school if they aren't offered a charter seat. The second most-likely counterfactual destiny for the younger applicant group is an innovation school, with nearly a third of non-offered compliers enrolling in one of these. The likelihood of an enrollment destiny outside the charter, traditional, and innovation sectors is much smaller.

[22] Innovation waivers are subject to approval by the Denver Classroom Teachers Association (which organizes Denver public school teachers' bargaining unit), and they allow, for example, increased instruction time. DPS innovation schools appear to have much in common with Boston's pilot schools, a model examined in Abdulkadiroğlu et al. (2011).

[23] Briefly, our procedure puts $(1-C_i)1(W_i = j)$ on the left hand side of a version of equation (8) with endogenous variable $1 - C_i$. The coefficient on this endogenous variable is an estimate of $E[1(W_{0i} = j)|C_{1i} > C_{0i}, X_i]$. The covariates and sample used here are the same as used to construct the 2SLS impact estimates reported in column 1 of Table 7.

Isolating an Innovation School Effect

The outsize role of innovation schools in counterfactual destinies motivates an empirical strategy that allows for distinct charter and innovation school treatment effects. By pulling innovation schools out of the non-charter counterfactual, we capture charter treatment effects driven mainly by the contrast between charter and traditional public schools. Comparisons with a more homogeneous counterfactual also mitigates bias that might arise from violations of the exclusion restriction (discussed in Section 3.4). And, of course, the innovation treatment effect is also of interest in its own right.

For the purposes of this discussion, we code the sector variable, W_i, as taking on the value 2 for innovation schools, the value 1 for charters, and 0 otherwise. The corresponding potential outcomes are Y_{2i}, Y_{1i}, and Y_{0i}. In principal, this leads to multiple heterogenous causal effects, $Y_{2i} - Y_{0i}$ and $Y_{1i} - Y_{0i}$, but the identification of multiple-treatment models with unrestricted heterogeneity raises issues that go beyond the scope of this paper.[24] Defining constant effects with the notation

$$\begin{aligned} Y_{2i} - Y_{0i} &= \beta_2, \\ Y_{1i} - Y_{0i} &= \beta_1, \end{aligned}$$

our two-sector identification strategy can be motivated by the conditional independence assumption,

$$Y_{0i} \perp\!\!\!\perp Z_i | \theta_i, \tag{11}$$

where Z_i is a categorical variable that records DA-generated offers in each sector sector (charter, innovation, other).

The instruments here are indicators for charter and innovation-sector offers, $D_i^1 = 1[Z_i = 1]$ and $D_i^2 = 1[Z_i = 2]$. These dummy instruments are used in a 2SLS procedure with two endogenous variables, C_i^1 for charter school enrollment and C_i^2 for innovation school enrollment. Propensity score conditioning is justified by the fact that conditional independence relation (11) implies

$$Y_{0i} \perp\!\!\!\perp Z_i \mid p_1(\theta), p_2(\theta), \tag{12}$$

where $p_1(\theta) = E[D_i^1|\theta]$ and $p_2(\theta) = E[D_i^2|\theta]$.

The 2SLS setup in this case consists of

$$Y_i = \sum_x \alpha_1(x) d_i^1(x) + \sum_x \alpha_2(x) d_i^2(x) + \beta_1 C_i^1 + \beta_2 C_i^2 + \epsilon_i, \tag{13}$$

$$C_i^1 = \sum_x \gamma_{11}(x) d_i^1(x) + \sum_x \gamma_{12}(x) d_i^2(x) + \delta_{11} D_i^1 + \delta_{12} D_i^2 + \nu_i, \tag{14}$$

$$C_i^2 = \sum_x \gamma_{21}(x) d_i^1(x) + \sum_x \gamma_{22}(x) d_i^2(x) + \delta_{21} D_i^1 + \delta_{22} D_i^2 + \eta_i, \tag{15}$$

[24] See Behaghel et al. (2013) and Blackwell (2015) for recent progress on this issue.

where the dummy control variables, $d_i^1(x)$ and $d_i^2(x)$, saturate estimates of the scores for each treatment, $\hat{p}_1(\theta_i)$ and $\hat{p}_2(\theta_i)$, with corresponding score effects denoted by γ's and α's in the first and second stage models. The sample used for this analysis is the union of charter and innovation school applicants.

As noted by Imbens (2000) and Yang et al. (2014), the key conditional independence relation in this context (equation (12)) suggests we should choose a parameterization that fixes conditional probabilities of assignment for all treatment levels *jointly*. Joint score control replaces the additive score controls in equations (13), (14), and (15) with score controls of the form

$$d_i^{12}(x^1, x^2) = 1[\hat{p}_1(\theta_i) = x^1, \hat{p}_2(\theta_i) = x^2],$$

where hats denote score estimates and the indices, x^1 and x^2, run independently over all values in the support for each score. This model generates far more score fixed effects than does equation (13).[25] Fortunately, however, the algebra of 2SLS obviates the need for joint score control; additive control is enough.

To see why additive control is adequate for 2SLS models that exploit (12), note first that 2SLS estimates of the second stage equation, (13), can be obtained by first regressing each offer dummy on the full set of $d_i^1(x)$ and $d_i^2(x)$ and then using the residuals from this regression as instruments after dropping these controls from the model (see, e.g., Angrist and Pischke 2009). Note also that both sets of regressors in this auxiliary first-step model are functions of type. A regression of D_i^j on the full set of $d_i^1(x)$ and $d_i^2(x)$ therefore returns fitted values given by $E[D^j|\theta] = p_j(\theta_i)$.

Suppose now that we replace additive controls, $d_i^1(x)$ and $d_i^2(x)$, with the full set of dummies, $d_i^{12}(x^1, x^2)$, parameterizing the jointly-controlled model. Since the model here is saturated, a regression of D_i^j on the full set of $d_i^{12}(x^1, x^2)$ dummies recovers the conditional expectation function of offers given both scores. By the law of iterated expectations, however, this is

$$E[D_i^j|p_1(\theta_i), p_2(\theta_i)] = E\{E[D_i^j|\theta_i]|p_1(\theta_i), p_2(\theta_i)]\} = p_j(\theta_i).$$

From this we conclude that the IV equivalent of 2SLS is the same in the additive and jointly-controlled models.[26]

As a benchmark, columns 1-2 of Table 10 compare charter-only and innovation-only estimates computed using DA (frequency) score controls. Each sample is limited to applicants to the relevant sector.[27] A parallel set of single-sector estimates using simulated score controls appears in columns 5 and 6. The innovation first stage (the effect of an innovation school offer on innovation school enrollment) is around 0.35. The pooled single-sector charter estimates in Table 10 are the same as those in Table 7. Not surprisingly in view of the substantially reduced

[25] When $p_1(\theta)$ takes on k_1 values and $p_2(\theta)$ takes on k_2 values, the additive model has k_1+k_2-2 score parameters, while the joint model has $k_1k_2 - 1$.

[26] This conclusion holds in the population, but need not hold exactly in our data (because scores here are estimated by something more elaborate than a sample mean conditional on type) or for models that include additional covariates beyond saturated score controls.

[27] Appendix Table B6 lists innovation schools and describes the random assignment pattern at these schools along the lines of Table 1 for charter schools. Covariate balance and differential attrition results for innovation schools are reported in Appendix Table B7.

number of applicants with non-trivial innovation offer risk (546 in column 2 and 613 in column 6 of Table 10), and the smaller innovation first stage, the innovation attendance effects are relatively imprecise. This imprecision notwithstanding, the innovation-only model generates a large negative and marginally significant effect on reading when estimated with the DA score.

2SLS estimates of equation (13) appear in columns 3 and 7 of Table 10. Charter school effects change little in this specification, but (insignificant) negative innovation estimates for math flip to positive when estimated using a model that also isolates charter treatment effects. The negative innovation school effects on reading seen in columns 2 and 6 also become smaller in the two-endogenous-variables models. Most interestingly, perhaps, the marginally significant positive charter school effect on reading (when estimating using DA score controls) also disappears. While charter students' reading performance exceeds what we can expect to see were these students to enroll in a mix of traditional and (low-performing) innovation schools, the reading gap between charters and traditional public schools is a little smaller.

As the theoretical discussion above leads us to expect, the results of estimation with joint score controls, shown in columns 4 and 8 of Table 10, differ little from the estimates constructed using additive score controls reported in columns 3 and 7 (a marginally significant though still imprecisely estimate positive innovation effect on math scores emerges in column 4). Overall, it seems fair to say that the findings on charter effectiveness in Table 7 stand when charter effects are estimated using a procedure that removes the innovation sector from the charter enrollment counterfactual.

5 Summary and Directions for Further Work

We investigate research strategies that exploit the random lottery number embedded in market design solutions to school matching problems. The most important fruit of this inquiry is the DA propensity score, an easily-computed formula for the conditional probability of assignment to particular schools as a function of student preferences and priorities. The DA propensity score maximizes the number of students and schools that can be studied by random assignment, while also revealing the nature of the experimental design generated as a by-product of deferred acceptance and related matching schemes. We also show how the DA score can be used to estimate causal effects in models with multiple sectors or schools.

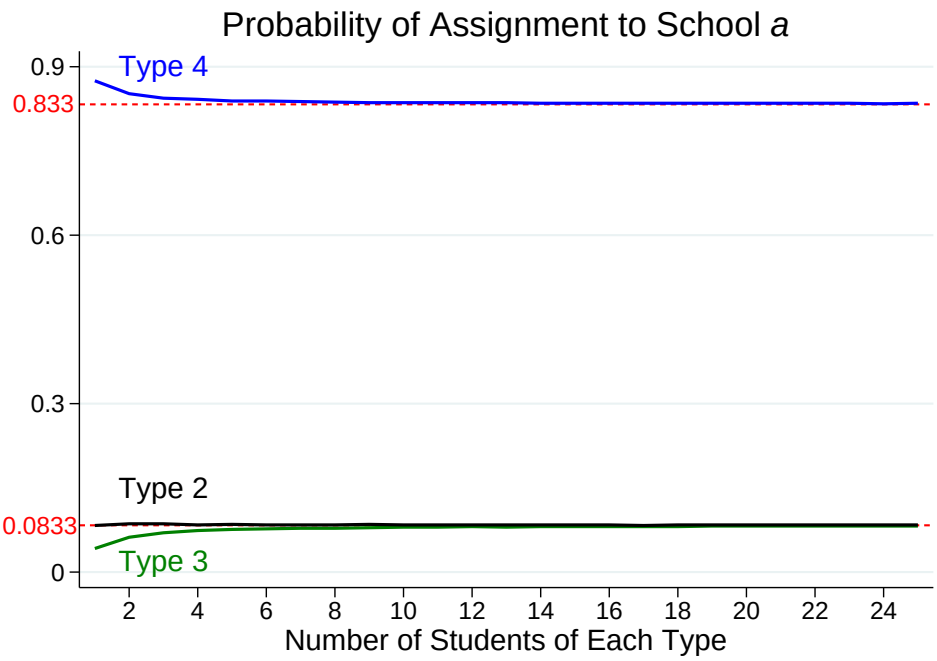
A score-based analysis of data from Denver's unified school match reveals substantial gains from attendance at one of Denver's many charter schools. The resulting charter effects are similar to those computed using single-school lottery strategies for Boston's charters reported in Abdulkadiroğlu et al. (2011). At the same time, as with previously reported results for Boston Pilot schools, Denver's Innovation model does not appear to generate substantial achievement gains. Our analysis focuses on defining and estimating the DA propensity score, giving less attention to the problem of how best to use the score for estimation. Still, simple 2SLS procedures seem to work well, and the resulting estimates of DPS charter effects differ little from those generated by semiparametric alternatives.

The methods developed here should be broadly applicable to markets using the DA family of mechanisms for centralized assignment. At the same time, some markets and matches use mech-

anisms not covered by the DA framework. Most important on this list is the top trading cycles (TTC) mechanism (Shapley and Scarf, 1974; Abdulkadiroğlu and Sönmez, 2003), which allows students to trade priorities rather than treating priorities as fixed. We hope to report theoretical results on the TTC propensity score soon, along with results from an application to New Orleans Recovery School District, which has experimented with TTC matching (Abdulkadiroğlu et al., 2014b).

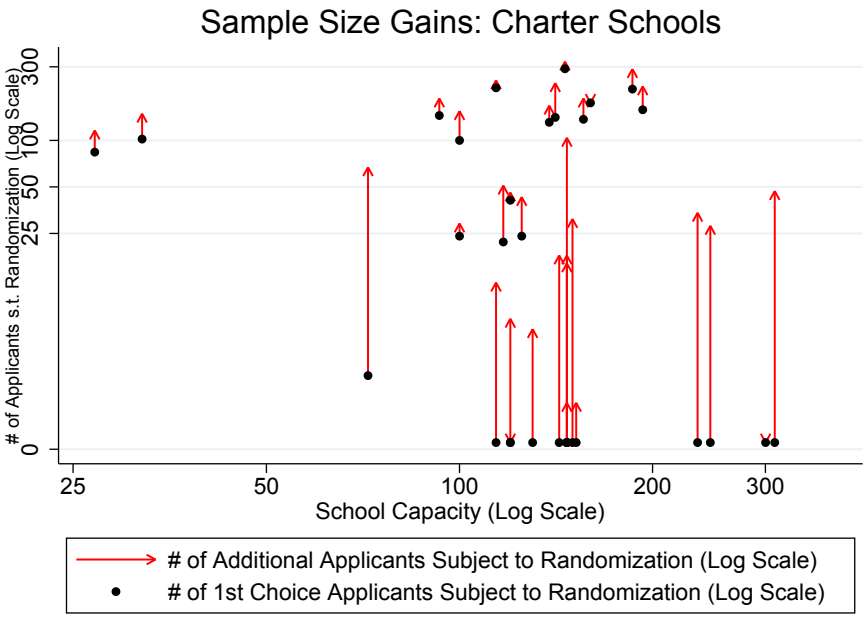
Finally, many matching problems, such as the selective exam schools analyzed by Jackson (2010); Dobbie and Fryer (2014); Abdulkadiroğlu et al. (2014a); Lucas and Mbiti (2014); Pop-Eleches and Urquiola (2013) use non-randomly-assigned tie-breakers rather than a lottery. These schemes embed a regression discontinuity design inside a market design rather than embedding a randomized trial. The question of how best to define and exploit the DA propensity score for markets that combine regression-discontinuity tie-breaking with market design matchmaking is a natural next step on the market-design-meets-research-design agenda.

Figure 1: Propensity Scores and Market Size in in Example 2

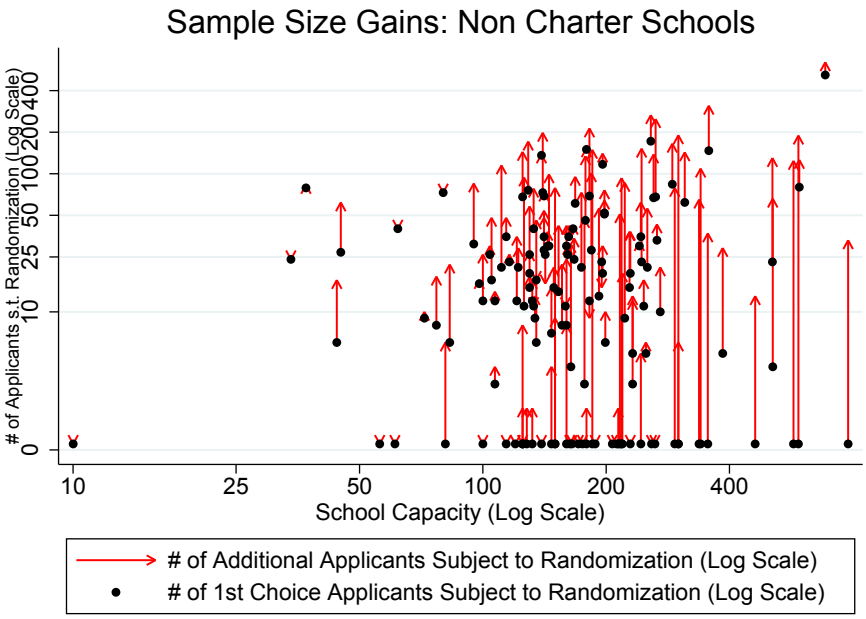


Notes: This figure plots finite-market propensity scores for expansions of Example 2 in Section 2.2. For each value of the x axis, we consider an expansion of the example with x students of each type. The propensity scores plotted here were computed by drawing lottery numbers 100,000 times.

Figure 2: Sample Size Gains from the Propensity Score Strategy



(a)



(b)

Notes: These figures compare the sample size under our DA propensity score strategy to that under the first choice strategy. Down arrows mean the two empirical strategies produce the same number of applicants subject to randomization at the corresponding schools. We say a student is subject to randomization at a school if the student has the DA propensity score (frequency) of assignment to that school that is neither 0 nor 1.

Table 1: DPS charter schools

School	Total applicants (1)	Applicants offered seats (2)	Propensity score in (0,1) DA score (frequency) (3)	DA score (formula) (4)	Simulated score (5)	Simulated score (first choice) (6)
Elementary and middle schools						
Cesar Chavez Academy Denver	62	9	7	9	8	3
Denver Language School	4	0	0	0	0	0
DSST: Cole	281	129	31	40	44	0
DSST: College View	299	130	47	67	68	0
DSST: Green Valley Ranch	1014	146	324	344	357	291
DSST: Stapleton	849	156	180	189	221	137
Girls Athletic Leadership School	221	86	18	40	48	0
Highline Academy Charter School	159	26	69	78	84	50
KIPP Montbello College Prep	211	39	36	48	55	20
KIPP Sunshine Peak Academy	389	83	41	42	44	36
Odyssey Charter Elementary	215	6	20	21	22	14
Omar D. Blair Charter School	385	114	135	141	182	99
Pioneer Charter School	25	5	0	2	2	0
SIMS Fayola International Academy Denver	86	37	7	18	20	0
SOAR at Green Valley Ranch	85	9	41	42	43	37
SOAR Oakland	40	4	0	9	7	2
STRIVE Prep - Federal	621	138	170	172	175	131
STRIVE Prep - GVR	324	112	104	116	118	0
STRIVE Prep - Highland	263	112	2	21	18	0
STRIVE Prep - Lake	320	126	18	26	26	0
STRIVE Prep - Montbello	188	37	16	31	35	0
STRIVE Prep - Westwood	535	141	235	238	239	141
Venture Prep	100	50	12	17	17	0
Wyatt Edison Charter Elementary	48	4	0	3	2	0
High schools						
DSST: Green Valley Ranch	806	173	290	343	330	263
DSST: Stapleton	522	27	116	117	139	96
Southwest Early College	265	76	34	47	55	0
Venture Prep	140	39	28	42	45	0
KIPP Denver Collegiate High School	268	60	29	37	40	24
SIMS Fayola International Academy Denver	71	15	6	22	22	0
STRIVE Prep - SMART	383	160	175	175	175	175

Notes: This table describes DPS charter applications. Column 1 reports the number of applicants ranking each school. Columns 3-6 count applicants with propensity score values strictly between zero and one according to different score computation methods. Column 6 shows the subset of applicants from column 5 who rank each school as their first choice.

Table 2: DA Score anatomy

				DA Score = 0			DA Score in (0,1)		DA Score = 1
				Θ_s^n	Θ_s^c	Θ_s^a	Θ_s^c	Θ_s^a	Θ_s^a
	Eligible applicants	Capacity	Offers	0≤MID≤1	MID ≥ τ_s	MID = 1	MID < τ_s	0<MID<1	MID = 0
Campus	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GVR	324	147	112	0	0	159	0	116	49
Lake	274	147	126	0	0	132	0	26	116
Highland	244	147	112	0	0	121	0	21	102
Montbello	188	147	37	0	0	128	0	31	29
Federal	574	138	138	78	284	3	171	1	37
Westwood	494	141	141	53	181	4	238	0	18

Notes: This table shows how formula scores are determined for STRIVE school seats in grade 6 (all 6th grade seats at these schools are assigned in a single bucket; ineligible applicants, who have a score of zero, are omitted). Column 3 records offers made to these applicants. Columns 4-6 show the number of applicants in partitions with a score of zero. Columns 7 and 8 show the number of applicants subject to random assignment. Column 9 shows the number of applicants with certain offers.

Table 3: DPS student characteristics

				Propensity score in (0,1)			
				DA score (frequency)		Simulated score	
	Denver students	SchoolChoice applicants	Charter applicants	Charter applicants	Charter students	Charter applicants	Charter students
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Origin school is charter	0.133	0.080	0.130	0.259	0.371	0.230	0.357
Female	0.495	0.502	0.518	0.488	0.496	0.506	0.511
Race							
Hispanic	0.594	0.593	0.633	0.667	0.713	0.636	0.711
Black	0.141	0.143	0.169	0.181	0.161	0.192	0.168
White	0.192	0.187	0.124	0.084	0.062	0.098	0.059
Asian	0.034	0.034	0.032	0.032	0.039	0.033	0.037
Gifted	0.171	0.213	0.192	0.159	0.152	0.165	0.149
Bilingual	0.039	0.026	0.033	0.038	0.042	0.032	0.037
Subsidized lunch	0.753	0.756	0.797	0.813	0.818	0.800	0.823
Limited English proficient	0.285	0.290	0.324	0.343	0.378	0.337	0.380
Special education	0.119	0.114	0.085	0.079	0.068	0.083	0.070
Baseline scores							
Math	0.000	0.015	0.021	0.037	0.089	0.037	0.062
Reading	0.000	0.016	0.005	-0.011	0.007	0.008	-0.002
Writing	0.000	0.010	0.006	0.001	0.039	0.016	0.035
N	40,143	10,898	4,964	1,436	828	1,523	781

Notes: This table decribes the population of Denver 3rd-9th graders in 2011-2012, the baseline and application year. Statistics in column 1 are for charter and non-charter students. Column 2 describes the subset that submitted an application to the SchoolChoice system for a seat in grades 4-10 at another DPS school in 2012-2013. Column 3 reports values for applicants ranking any charter school. Columns 4-7 show statistics for charter applicants with propensity score values strictly between zero and one. Test scores are standardized to the population in column 1.

Table 4: Expected balance

			Propensity score controls					
			DA score (frequency)			Simulated score		
				Nonparametric			Nonparametric	
Covariate	Non-offered mean (1)	No controls (2)	Linear control (3)	Rounded (hundredths) (4)	Saturated (5)	Linear control (6)	Rounded (hundredths) (7)	Rounded (ten thousandths (8)
				A. Application covariates				
Number of schools ranked	4.375	-0.341	0.107	0.065	0.052	0.012	0.013	0.001
Number of charter schools ranked	1.426	0.474	0.109	0.069	0.055	0.004	0.004	0.000
First school ranked is charter	0.290	0.616	0.004	0.004	0.001	-0.002	-0.002	0.000
				B. Baseline covariates				
Origin school is charter	0.083	0.115	-0.017	-0.002	0.001	0.001	-0.001	0.000
Female	0.521	-0.007	0.002	0.003	0.003	0.002	0.001	0.001
Race								
Hispanic	0.595	0.094	-0.010	-0.005	-0.005	0.003	0.002	0.000
Black	0.182	-0.031	0.005	0.001	0.001	-0.001	0.000	0.000
Gifted	0.201	-0.022	-0.003	-0.004	-0.005	0.000	0.001	0.000
Bilingual	0.025	0.020	0.001	0.001	0.001	0.000	0.001	0.000
Subsidized lunch	0.767	0.073	0.003	0.003	0.003	0.002	0.002	0.000
Limited English proficient	0.290	0.084	-0.004	-0.004	-0.005	0.000	0.000	-0.001
Special education	0.087	-0.004	-0.004	-0.004	-0.003	0.001	0.000	0.000
Baseline scores								
Math	0.017	0.010	-0.023	-0.019	-0.020	-0.001	-0.002	-0.002
Reading	0.034	-0.070	-0.016	-0.014	-0.014	-0.002	-0.002	-0.003
Writing	0.029	-0.056	-0.019	-0.016	-0.016	-0.001	-0.002	-0.002
Average risk set points of support			87	39	47	1,148	51	121

Notes: This table reports average covariate balance by charter offer status across 400 lottery draws, with DA rerun each time. Balance is estimated by regressing each covariate on an any-charter simulated offer dummy, controlling for the propensity score variables indicated in each column header. The table reports averages of these balance coefficients. The sample includes applicants for 2012-13 charter seats in grades 4-10 who were enrolled in Denver at baseline. The charter offer variable indicates an offer at any charter school, excluding alternative charters. Column 1 reports the baseline characteristics of charter applicants who did not receive a charter offer. The average risk set points of support reported at the bottom of the table count the average number of unique values found in the support of the relevant propensity score. Except for columns 3 and 6, this excludes values of zero and one. The estimates in columns 4, 7 and 8 use score values rounded as indicated in the column header; the estimates in column 5 control for every score value seen in the data.

Table 5a: Statistical tests for balance in application covariates

		Propensity score controls						
		DA score (frequency)			Simulated score			
			Nonparametric			Nonparametric		
Application variable	No controls (1)	Linear control (2)	Rounded (hundredths) (3)	Saturated (4)	Linear control (5)	Rounded (hundredths) (6)	Rounded (ten thousandths) (7)	Full applicant type controls (8)
Number of schools ranked	-0.341***	0.097	0.059	0.028	0.014	0.001	-0.061	-0.015
	(0.046)	(0.103)	(0.095)	(0.094)	(0.102)	(0.095)	(0.125)	(0.042)
Number of charter schools ranked	0.476***	0.143***	0.100**	0.074	0.020	-0.017	0.009	0.007
	(0.024)	(0.052)	(0.047)	(0.047)	(0.048)	(0.043)	(0.061)	(0.010)
First school ranked is charter	0.612***	0.012	0.002	-0.001	-0.030	-0.042*	0.012	0.000
	(0.011)	(0.025)	(0.022)	(0.020)	(0.027)	(0.022)	(0.027)	(0.000)
N	4,964	1,436	1,289	1,247	1,523	1,290	681	301
Risk set points of support		88	40	47	1,148	51	126	61
Robust F-test for joint significance	1190	2.70	1.70	1.09	0.49	1.26	0.31	0.34
p-value	0.000	0.044	0.165	0.352	0.688	0.287	0.817	0.710

Notes: This table reports coefficients from regressions of the application variables in each row on a dummy for charter offers. The sample includes applicants for 2012-13 charter seats in grades 4-10 who were enrolled in Denver at baseline. Columns 1-7 are from regressions like those used to construct expected balance in Table 4, except that the tests reported here use realized DA offers, with test statistics and standard errors computed in the usual way. Column 8 reports the balance test generated by a regression with saturated controls for applicant type (that is, unique combinations of applicant preferences over school programs and school priorities in those programs). Robust standard errors are reported in parentheses. P-values for robust joint significance tests are estimated by stacking outcomes and clustering at the student level.

*significant at 10%; **significant at 5%; ***significant at 1%

Table 5b: Statistical tests for balance in student characteristics

		Propensity score controls					
		DA score (frequency)			Simulated score		
			Nonparametric			Nonparametric	
	No controls	Linear control	Rounded (hundredths)	Saturated	Linear control	Rounded (hundredths)	Rounded (ten thousandths)
Student characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Origin school is charter	0.108***	-0.051**	-0.037**	-0.029*	-0.039*	-0.036**	-0.037*
	(0.010)	(0.024)	(0.017)	(0.017)	(0.023)	(0.017)	(0.022)
Female	-0.005	0.024	0.021	0.019	0.016	0.030	0.010
	(0.014)	(0.034)	(0.034)	(0.034)	(0.033)	(0.034)	(0.054)
Race							
Hispanic	0.095***	-0.022	-0.013	-0.007	0.005	-0.001	-0.018
	(0.014)	(0.031)	(0.028)	(0.028)	(0.031)	(0.029)	(0.042)
Black	-0.033***	-0.002	-0.005	-0.007	-0.012	-0.012	0.011
	(0.011)	(0.026)	(0.025)	(0.025)	(0.026)	(0.026)	(0.039)
Gifted	-0.028**	-0.026	-0.028	-0.030	-0.032	-0.035	-0.037
	(0.011)	(0.026)	(0.026)	(0.026)	(0.025)	(0.026)	(0.042)
Bilingual	0.023***	0.016	0.014	0.015	0.012	0.014	0.011
	(0.005)	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)	(0.021)
Subsidized lunch	0.073***	-0.003	-0.004	0.001	0.001	-0.005	0.024
	(0.011)	(0.027)	(0.025)	(0.025)	(0.027)	(0.026)	(0.037)
Limited English proficient	0.086***	-0.002	-0.002	0.001	0.011	0.001	0.004
	(0.014)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.053)
Special education	0.004	0.034**	0.032*	0.032*	0.043**	0.044**	0.035
	(0.008)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.028)
N	4,964	1,436	1,289	1,247	1,523	1,290	681
Baseline scores							
Math	-0.002	-0.087	-0.083	-0.082	-0.068	-0.078	-0.053
	(0.027)	(0.061)	(0.060)	(0.061)	(0.061)	(0.061)	(0.094)
Reading	-0.085***	-0.096*	-0.100*	-0.108*	-0.081	-0.086	-0.070
	(0.026)	(0.057)	(0.056)	(0.056)	(0.056)	(0.056)	(0.087)
Writing	-0.072***	-0.097*	-0.096*	-0.101*	-0.085	-0.094*	-0.053
	(0.026)	(0.056)	(0.054)	(0.055)	(0.055)	(0.054)	(0.083)
N	4,889	1,420	1,275	1,234	1,504	1,275	675
Robust F-test for joint significance	19.1	1.20	1.13	0.99	1.04	1.35	0.71
p-value	0.000	0.278	0.329	0.454	0.408	0.183	0.743

See notes to Table 5a.

*significant at 10%; **significant at 5%; ***significant at 1%

Table 6: Comparison of 2SLS and semiparametric estimates of charter effects

	Frequency (saturated)		Formula (saturated)		Simulation (hundredths)	
	2SLS	Semiparametric	2SLS	Semiparametric	2SLS	Semiparametric
	(1)	(2)	(3)	(4)	(5)	(6)
Math	0.496***	0.443***	0.524***	0.486***	0.543***	0.474**
	(0.076)	(0.105)	(0.071)	(0.105)	(0.075)	(0.212)
	{0.071}		{0.076}		{0.079}	
Reading	0.127*	0.106	0.120*	0.118	0.106	0.127
	(0.065)	(0.107)	(0.073)	(0.115)	(0.069)	(0.173)
	{0.065}		{0.069}		{0.071}	
Writing	0.325***	0.326***	0.356***	0.364***	0.324***	0.305**
	(0.079)	(0.102)	(0.082)	(0.113)	(0.079)	(0.145)
	{0.077}		{0.080}		{0.080}	
N	1,102	1,093	1,083	1,081	1,137	1,137

Notes: This table compares 2SLS and semiparametric estimates of charter attendance effects on the 2012-13 TCAP scores of Denver 4th-10th graders. The instrument is an any-charter offer dummy. The semiparametric estimator is described in Section 3.5. In addition to score variables, 2SLS estimates include controls for grade tested, gender, origin school charter status, race, gifted status, bilingual status, subsidized price lunch eligibility, special education, limited English proficient status, and baseline test scores. Semiparametric models use these same variables as controls when computing the score weighting function. Standard errors in parentheses are from a Bayesian bootstrap. Conventional robust standard errors for 2SLS estimates are reported in braces.

*significant at 10%; **significant at 5%; ***significant at 1%

Table 7: Comparison of 2SLS and OLS estimates of charter attendance effects

	2SLS estimates					
	DA score					
	Frequency (saturated) (1)	Formula (saturated) (2)	Simulated score rounded (hundredths) (3)	No score controls (4)	OLS (5)	OLS with score controls (6)
First stage	0.410***	0.389***	0.377***	0.683***		
	(0.031)	(0.032)	(0.032)	(0.012)		
Math	0.496***	0.524***	0.543***	0.306***	0.304***	0.386***
	(0.071)	(0.076)	(0.079)	(0.021)	(0.015)	(0.034)
Reading	0.127**	0.120*	0.106	0.093***	0.103***	0.093***
	(0.065)	(0.069)	(0.071)	(0.020)	(0.014)	(0.029)
Writing	0.325***	0.356***	0.324***	0.183***	0.180***	0.202***
	(0.077)	(0.080)	(0.080)	(0.023)	(0.015)	(0.036)
N	1,102	1,083	1,137	4,317	4,317	1,102

Notes: This table compares 2SLS and OLS estimates of charter attendance effects using the same sample and instruments as for Table 6. The OLS estimates in column 6 are from a model that includes saturated control for frequency estimates of the DA score. In addition to score variables, all models include controls for grade tested, gender, origin school charter status, race, gifted status, bilingual status, subsidized price lunch eligibility, special education, limited English proficient status, and baseline test scores. Robust standard errors are reported in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 8: Other IV strategies

	Charter attendance effect				
	Offer instrument with DA score (frequency) controls (saturated)	First choice charter offer with risk set controls	Qualification instrument with risk set controls	Sample size increase for equivalent gain (col 2 vs col 1)	Sample size increase for equivalent gain (col 3 vs col 1)
	(1)	(2)	(3)	(4)	(5)
		A. First stage estimates			
First stage for charter offers	1.000	0.774***	0.476***		
	--	(0.026)	(0.024)		
First stage for charter enrollment	0.410***	0.323***	0.178***		
	(0.031)	(0.035)	(0.027)		
		B. 2SLS estimates			
Math	0.496***	0.596***	0.409***		
	(0.071)	(0.102)	(0.149)	2.0	4.4
Reading	0.127**	0.227**	0.229		
	(0.065)	(0.102)	(0.144)	2.5	4.9
Writing	0.325***	0.333***	0.505***		
	(0.077)	(0.119)	(0.162)	2.4	4.5
N (students)	1,102	1,125	1,969		
N (schools)	30	15	24		

Notes: This table compares alternative 2SLS estimates of charter attendance effects using the same sample and control variables used to construct the estimates in Tables 6-7. Column 1 repeats the estimates using a DA offer instrument from column 1 in Table 7. The row labeled "First stage for charter offers" reports the coefficient from a regression of any-charter offer dummy (the instrument used in column 1) on other instruments, conditioning on the same controls used in the corresponding first stage estimates for charter enrollment. Column 2 reports 2SLS estimates computed using a first-choice charter offer instrument. Column 3 reports charter attendance effects computed using an any-charter qualification instrument. These alternative IV models control for risk sets making the first-choice and qualification instruments conditionally random; see Section 4.5 for details. Columns 4 and 5 report the multiples of the first-choice offer sample size and qualification sample size needed to achieve a precision gain equivalent to the gain from using the any-charter offer instrument. The last row counts the number of schools for which we observe in-sample variation in offer rates conditional on the score controls included in the model.

*significant at 10%; **significant at 5%; ***significant at 1%

Table 9: Enrollment destinies for charter applicants

	All charter applicants		Charter applicants with DA score (frequency) in (0,1)			
			All applicants		Compliers	
	No charter offer	Charter offer	Non-offered mean	First stage + col 3	No charter offer	Charter offer
	(1)	(2)	(3)	(4)	(5)	(6)
Enrolled in a study charter	0.147	0.865	0.347	0.757	--	1.000
… in a traditional public	0.405	0.081	0.257	0.054	0.497	--
… in an innovation school	0.234	0.023	0.241	0.107	0.328	--
… in a magnet school	0.192	0.021	0.116	0.078	0.094	--
… in an alternative school	0.009	0.005	0.018	0.006	0.030	--
… in a contract school	0.012	0.004	0.018	-0.001	0.047	--
… in a non-study charter	0.001	0.000	0.002	0.000	0.005	--
N	2,555	1,833	498	1,102	--	--

Notes: This table describes school enrollment outcomes for charter applicants in the sample used to construct the estimates reported in Table 7. Columns 1-2 show enrollment by sesctor for all applicants without and with a charter offer. The remaining columns look only at those with a DA (frequency) score strictly between zero and one. Column 4 adds the non-offered mean in column 3 to the first stage estimate of the effect of charter offers on charter enrollment. School sectors are classified by grade. Innovation schools design and implement innovative practices to improve student outcomes. Magnet schools serve students with particular styles of learning. Alternative schools serve students struggling with academics, behavior, attendance, or other factors that may prevent them from succeeding in a traditional school environment; the latter offer faster pathways toward high school graduation, such as GED preparation and technical education. There is a single contract school, Escuela Tlatelolco, a private school contracted to serve DPS students, and a single non-study charter that closed in May 2013. Complier means in columns 5 and 6 were estimated using the 2SLS procedures described by Abadie(2002), with the same propensity score and covariate controls as were used to construct the estimates in Table 7.

Table 10: DPS charter and innovation school attendance effects

	DA score (frequency) controls (saturated)				Simulated score controls rounded (hundredths)			
			Charter and innovation				Charter and innovation	
	Charter only (1)	Innovation only (2)	Additive score controls (3)	Joint score controls (4)	Charter only (5)	Innovation only (6)	Additive score controls (7)	Joint score controls (8)
Charter First Stage	0.410***	--	0.405***	0.398***	0.377***	--	0.437***	0.417***
	(0.031)	--	(0.034)	(0.035)	(0.032)	--	(0.032)	(0.035)
Innovation First Stage	--	0.348***	0.347***	0.348***	--	0.345***	0.301***	0.300***
	--	(0.042)	(0.042)	(0.044)	--	(0.041)	(0.040)	(0.043)
				A. Math				
Charter	0.496***	--	0.534***	0.517***	0.543***	--	0.618***	0.550***
	(0.071)	--	(0.077)	(0.082)	(0.079)	--	(0.073)	(0.082)
Innovation	--	-0.035	0.177	0.286*	--	-0.180	0.199	0.146
	--	(0.136)	(0.134)	(0.147)	--	(0.137)	(0.159)	(0.174)
				B. Reading				
Charter	0.127**	--	0.076	0.072	0.106	--	0.105	0.089
	(0.065)	--	(0.078)	(0.084)	(0.071)	--	(0.075)	(0.085)
Innovation	--	-0.285**	-0.231	-0.190	--	-0.203	-0.074	-0.162
	--	(0.141)	(0.153)	(0.165)	--	(0.136)	(0.161)	(0.185)
				C. Writing				
Charter	0.325***	--	0.357***	0.334***	0.324***	--	0.348***	0.393***
	(0.077)	--	(0.087)	(0.094)	(0.080)	--	(0.079)	(0.087)
Innovation	--	-0.119	0.115	0.052	--	-0.057	0.063	0.004
	--	(0.136)	(0.148)	(0.156)	--	(0.132)	(0.153)	(0.167)
N	1,102	546	1,418	1,274	1,137	613	1,583	1,274

Notes: This table reports 2SLS estimates of charter and innovation attendance effects for applicants to schools in one or both sectors. The estimates for charter applicants in columns 1 and 5 are the same as reported in column 1 of Table 7. Columns 2 and 6 report innovation attendance effects for innovation applicants, estimated in models using an innovation offer instrument and innovation-specific saturated score controls constructed like those used for charter applicants. Columns 3 and 7 report coefficients from a two-endogenous-variable/two-instrument 2SLS model for the attendance effects of charters and innovations, conditioning additively on charter-specific and innovation-specific saturated score controls. Columns 4 and 8 show results from joint-effect models that add interactions between the two scores to the specification that generated column 7.
*significant at 10%; **significant at 5%; ***significant at 1%

A Theoretical Appendix

A.1 Defining DA: Details

Our general formulation defines the DA match as determined by cutoffs found in the limit of a sequence. Recall that these cutoffs evolve according to

$$c_s^{t+1} = \begin{cases} K+1 & \text{if } F(Q_s(\mathbf{c}^t)) < q_s, \\ \max\left\{x \in [0, K+1] \mid F(\{i \in Q_s(\mathbf{c}^t) \text{ such that } \pi_{is} \leq x\}) \leq q_s\right\} & \text{otherwise,} \end{cases}$$

where $Q_s(\mathbf{c}^t)$ is the demand for seats at school s for a given vector of cutoffs $\mathbf{c}^t$ and is defined as

$$Q_s(\mathbf{c}^t) = \{i \in I \mid \pi_{is} \leq c_s^t \text{ and } s \succ_i \tilde{s} \text{ for all } \tilde{s} \in S \text{ such that } \pi_{i\tilde{s}} \leq c_{\tilde{s}}^t\}. \tag{16}$$

The following result confirms that these limiting cutoffs exist, i.e., that the sequence $\mathbf{c}^t$ converges.

Proposition 1. *Consider an economy described by a distribution of students F and school capacities as defined in Section 3.1. Construct a sequence of cutoffs, c_s^t, for this economy as described above. Then, $\lim_{t\to\infty} c_s^t$ exists.*

Proof. c_s^t is well-defined for all $t \geq 1$ and all $s \in S$ since it is either $K+1$ or the maximizer of a continuous function over a compact set. We will show by induction that $\{c_s^t\}$ is a decreasing sequence for all s.

For the base case, $c_s^2 \leq c_s^1$ for all s since $c_s^1 = K+1$ and $c_s^2 \leq K+1$ by construction.

For the inductive step, suppose that $c_s^t \leq c_s^{t-1}$ for all s and all $t = 1, ..., T$. For each s, if $c_s^T = K+1$, then $c_s^{T+1} \leq c_s^T$ since $c_s^t \leq K+1$ for all t by construction. Otherwise, suppose to the contrary that $c_s^{T+1} > c_s^T$. Since $c_s^T < K+1$, $F(\{i \in Q_s(\mathbf{c}^{T-1}) \text{ such that } \pi_{is} \leq c_s^T\}) = q_s$. Then,

$$\begin{aligned} &F(\{i \in Q_s(\mathbf{c}^T) \text{ such that } \pi_{is} \leq c_s^{T+1}\}) \\ &\quad = F(\{i \in Q_s(\mathbf{c}^T) \text{ such that } \pi_{is} \leq c_s^T\}) + F(\{i \in Q_s(\mathbf{c}^T) \text{ such that } c_s^T < \pi_{is} \leq c_s^{T+1}\}) \\ &\quad \geq F(\{i \in Q_s(\mathbf{c}^{T-1}) \text{ such that } \pi_{is} \leq c_s^T\}) + F(\{i \in Q_s(\mathbf{c}^T) \text{ such that } c_s^T < \pi_{is} \leq c_s^{T+1}\}) \quad (17) \\ &\quad \geq q_s + F(\{i \in Q_s(\mathbf{c}^T) \text{ such that } c_s^T < \pi_{is} \leq c_s^{T+1}\}) \quad (18) \\ &\quad > q_s. \quad (19) \end{aligned}$$

Expression (17) follows because

$$\begin{aligned} &\{i \in Q_s(\mathbf{c}^T) \text{ such that } \pi_{is} \leq c_s^T\} \\ &\qquad = \{i \in I \mid \pi_{is} \leq c_s^T \text{ and } s \succ_i \tilde{s} \text{ for all } \tilde{s} \in S \text{ such that } \pi_{i\tilde{s}} \leq c_{\tilde{s}}^T\} \\ &\qquad \supseteq \{i \in I \mid \pi_{is} \leq c_s^T \text{ and } s \succ_i \tilde{s} \text{ for all } \tilde{s} \in S \text{ such that } \pi_{i\tilde{s}} \leq c_{\tilde{s}}^{T-1}\} \quad (\text{by } c_{\tilde{s}}^T \leq c_{\tilde{s}}^{T-1}) \\ &\qquad = \{i \in Q_s(\mathbf{c}^{T-1}) \text{ such that } \pi_{is} \leq c_s^T\}. \end{aligned}$$

Expression (18) follows by the inductive assumption and since $c_s^T < K+1$.
Expression (19) follows since if $F(\{i \in Q_s(\mathbf{c}^T) \text{ such that } c_s^T < \pi_{is} \leq c_s^{T+1}\}) = 0$, then

$$F(\{i \in Q_s(\mathbf{c}^{T-1}) \text{ such that } \pi_{is} \leq c_s^{T+1}\}) = F(\{i \in Q_s(\mathbf{c}^{T-1}) \text{ such that } \pi_{is} \leq c_s^T\}) \leq q_s,$$

while $c_s^{T+1} > c_s^T$, contradicting the definition of c_s^T.

Expression (19) contradicts the definition of c^{T+1} since the cutoff at step $T+1$ results in an allocation that exceeds the capacity of school s. This therefore establishes the inductive step that $c_s^{T+1} \leq c_s^T$.

To complete the proof of the proposition, observe that since $\{c_s^t\}$ is a decreasing sequence in the compact interval $[0, K+1]$, c_s^t converges by the monotone convergence theorem. □

Note that this result applies to the cutoffs for both finite and continuum economies. In finite markets, at convergence, these cutoffs produce the allocation we get from the usual definition of DA (e.g., as in Gale and Shapley (1962)). This can be seen by noting that

$$\begin{aligned}&\max\{x \in [0, K+1] \mid F(\{i \in Q_s(\mathbf{c}^t) \text{ such that } \pi_{is} \leq x\}) \leq q_s\} \\ &= \max\{x \in [0, K+1] \mid |\{j \in Q_s(\mathbf{c}^t) : \pi_{js} \leq x\}| \leq k_s\},\end{aligned}$$

implying that the tentative cutoff at school s in step t in our DA formulation, which is determined by the left hand side of this equality, is the same as that in Gale and Shapley (1962)'s DA formulation, which is determined by the right hand side of the equality. Our DA formulation and the Gale and Shapley (1962) formulation therefore produce the same cutoff at each step. This also implies that, in finite markets, our DA cutoffs are found in a finite number of iterations, since DA as described by Gale and Shapley (1962) converges in a finite number of steps.

A.2 Conditional Independence of DA-generated Offers

Proposition 2. *Let W_i be any variable that is independent of lottery numbers and write $P[D_i(s) = 1|W_i, \theta]$ for $P[D_i(s) = 1|W_i, \theta_i = \theta]$. Then*

$$P[D_i(s) = 1|W_i, \theta] = P[D_i(s) = 1|\theta].$$

Proof. Suppose that DA converges at iteration T. Demand given cutoffs, $Q(c^T)$, a function of random numbers, preferences, and priorities, determines the distribution of offers. Equation (16) therefore implies that

$$P[D_i(s) = 1|W_i, c^T, \theta] = P[D_i(s) = 1|c^T, \theta]. \tag{20}$$

Equation (20) does not contradict the fact that c^T is determined in part by interactions between types and realized lottery numbers, interactions that may distort the distribution of lottery numbers conditional on cutoffs. In particular, (20) holds even if conditioning on c^T makes the lottery number distribution depend on θ. This follows from the fact that, as a consequence of the definition of demand for DA, offers and cutoffs are *jointly* independent of W_i given θ.[28]

[28] Using the shorthand notation $P[D_i(s), W_i, c^T, \theta]$ to denote joint probability statements and the associated conditionals without specifying realized values, we have:

$$P[D_i(s)|W_i, c^T, \theta] = \frac{P[D_i(s), c^T|W_i, \theta]}{P[c^T|W_i, \theta]} = \frac{P[D_i(s), c^T|\theta]}{P[c^T|\theta]} = P[D_i(s)|c^T, \theta].$$

Joint independence is used for the second equality.

Note also that cutoffs satisfy

$$c^T \perp\!\!\!\perp W_i|\theta, \tag{21}$$

a consequence of the fact that c^T is the same for all i in every lottery draw.

Finally, using (20), we have

$$P[D_i(s)=1|W_i,\theta] = E\{E[D_i|W_i,\theta,c^T]|W_i,\theta\} = E\{E[D_i|c^T,\theta]|W_i,\theta\},$$

and the further implication by (21) that

$$E\{E[D_i|c^T,\theta]|W_i,\theta\} = E\{E[D_i|c^T,\theta]|\theta\} = P[D_i(s)=1|\theta],$$

completing the proof for finite markets.

Extension to the continuum follows from the definition of $Q(c)$ for the limiting allocation in the continuum and the fact that the limiting cutoff in the continuum, c, is non-stochastic. □

A.3 Proof of Theorem 1

Admissions cutoffs $\mathbf{c}$ in a continuum economy are invariant to lottery outcomes (r_i): DA in the continuum depends on (r_i) only through $F(I_0)$ for sets $I_0 = \{i \in I \mid \theta_i \in \Theta_0\}$ with various choices of Θ_0. In particular, $F(I_0)$ doesn't depend on lottery realizations. Likewise, marginal priority $\rho_{\tilde{s}}$ is uniquely determined for every school $\tilde{s}$.

Consider the propensity score for school s. Students who don't rank s have $\varphi_s(\theta) = 0$. Among those who do rank s, those of type $\theta \in \Theta_s^n$ have $\rho_{\theta s} > \rho_s$. Therefore $\varphi_s(\theta) = 0$ for every $\theta \in \Theta_s^n \cup (\Theta \backslash \Theta_s)$.

Students of type $\theta \in \Theta_s^a \cup \Theta_s^c$ may be assigned $\tilde{s} \in B_{\theta s}$, where $\rho_{\theta\tilde{s}} = \rho_{\tilde{s}}$. Since lottery numbers are uniform, the proportion of type θ students assigned some $\tilde{s} \in B_{\theta s}$ where $\rho_{\theta\tilde{s}} = \rho_{\tilde{s}}$ is $MID_{\theta\tilde{s}}$. In other words, the probability of not being assigned any $\tilde{s} \in B_{\theta s}$ where $\rho_{\theta\tilde{s}} = \rho_{\tilde{s}}$ for a type θ student is $1 - MID_{\theta s}$. Every student of type $\theta \in \Theta_s^a$ who is not assigned a higher choice is assigned s because $\rho_{\theta s} < \rho_s$, and so

$$\varphi_s(\theta) = (1 - MID_{\theta s}) \text{ for all } \theta \in \Theta_s^a.$$

Finally, consider students of type $\theta \in \Theta_s^c$ who are not assigned a higher choice. The fraction of students $\theta \in \Theta_s^c$ who are not assigned a higher choice is $1 - MID_{\theta s}$. Also, the random numbers of these students is larger than $MID_{\theta s}$. If $\tau_s < MID_{\theta s}$, then no such student is assigned s. If $\tau_s \geq MID_{\theta s}$, then the ratio of students that are assigned s within this set is given by $\frac{\tau_s - MID_{\theta s}}{1 - MID_{\theta s}}$. Hence, conditional on $\theta \in \Theta_s^c$ and not being assigned a choice higher than s, the probability of being assigned s is given by $\max\{0, \frac{\tau_s - MID_{\theta s}}{1 - MID_{\theta s}}\}$. Therefore,

$$\varphi_s(\theta) = (1 - MID_{\theta s}) \times \max\left\{0, \frac{\tau_s - MID_{\theta s}}{1 - MID_{\theta s}}\right\} \text{ for all } \theta \in \Theta_s^c.$$

A.4 Proof of Theorem 2

We complete the proof of Theorem 2 in Section 3.3 by proving the following two intermediate results.

Lemma 1. *(Cutoff almost sure convergence)* $\hat{\mathbf{c}}_n \xrightarrow{a.s.} \mathbf{c}$.

Lemma 2. *(Propensity score almost sure convergence) For all* $\theta \in \Theta$ *and* $s \in S$, $p_{ns}(\theta) \xrightarrow{a.s.} \varphi_s(\theta)$.

A.4.1 Proof of Lemma 1

We use the Extended Continuous Mapping Theorem (Theorem 19.1 in van der Vaart (2000)) to prove the lemma. We first show deterministic convergence of cutoffs in order to verify the assumptions of the theorem.

Modify the definition of F to describe the distribution of lottery numbers as well types: For any set of student types $\Theta_0 \subset \Theta$ and for any numbers $r_0, r_1 \in [0,1]$ with $r_0 < r_1$, define the set of students of types in Θ_0 with random numbers worse than r_0 and better than r_1 as

$$I(\Theta_0, r_0, r_1) = \{i \in I \mid \theta_i \in \Theta_0, r_0 < r_i \leq r_1\}.$$

In a continuum economy,

$$F(I(\Theta_0, r_0, r_1)) = E[1\{\theta_i \in \Theta_0\}] \times (r_1 - r_0),$$

where the expectation is assumed to exist. In a finite economy with n students,

$$F(I(\Theta_0, r_0, r_1)) = \frac{|I(\Theta_0, r_0, r_1)|}{n}.$$

Let $\mathcal{F}$ be the set of possible F's defined above. For any two distributions F and F', the supnorm metric is defined by

$$d(F, F') = \sup_{\Theta_0 \subset \Theta, r_0, r_1 \in [0,1]} |F(I(\Theta_0, r_0, r_1)) - F'(I(\Theta_0, r_0, r_1))|.$$

The notation is otherwise as in the text.

Proof. Consider a deterministic sequence of economies described by a sequence of distributions $\{f_n\}$ over students, together with associated school capacities, so that for all n, $f_n \in \mathcal{F}$ is a potential realization produced by randomly drawing n students and their lottery numbers from F. Assume that $f_n \to F$ in metric space $(\mathcal{F}, d)$. Let $\mathbf{c}_n$ denote the admissions cutoffs in f_n. Note the $\mathbf{c}_n$ is constant because this is the cutoff for a particular realized economy f_n.

The proof first shows deterministic convergence of cutoffs for any convergent subsequence of f_n. Let $\{\tilde{f}_n\}$ be a subsequence of realized economies $\{f_n\}$. The corresponding cutoffs are denoted $\{\tilde{\mathbf{c}}_n\}$. Let $\tilde{\mathbf{c}} \equiv (\tilde{c}_s)$ be the limit of $\tilde{\mathbf{c}}_n$. The following two claims establish that $\tilde{\mathbf{c}}_n \to \mathbf{c}$, the cutoff associated with F.

Claim 1. $\tilde{c}_s \geq c_s$ *for every* $s \in S$.

Proof of Claim 1. This is proved by contradiction in 3 steps. Suppose to the contrary that $\tilde{c}_s < c_s$ for some s. Let $S' \subset S$ be the set of schools the cutoffs of which are strictly lower under $\tilde{\mathbf{c}}$. For any $s \in S'$, define $I_n^s = \{i \in I | \tilde{c}_{ns} < \pi_{is} \leq c_s \text{ and } i \text{ ranks } s \text{ first}\}$ where I is the set of students in F, which contains the set of students in f_n for all n. In other words, I_n^s are the set of students ranking school s first who have a student rank in between $\tilde{c}_{ns}$ and c_s.

Step (a): We first show that for our subsequence, when the market is large enough, there must be some students who are in I_n^s. That is, there exists N such that for any $n > N$, we have $\tilde{f}_n(I_n^s) > 0$ for all $s \in S'$.

To see this, we begin by showing that for all $s \in S'$, there exists N such that for any $n > N$, we have $F(I_n^s) > 0$. Suppose, to the contrary, that there exists $s \in S'$ such that for all N, there exists $n > N$ such that $F(I_n^s) = 0$. When we consider the subsequence of realized economies $\{\tilde{f}_n\}$, we find that

$$\begin{aligned}
&\tilde{f}_n(\{i \in Q_s(\mathbf{c}_n) \text{ such that } \pi_{is} \leq c_s\}) \\
&\quad = \tilde{f}_n(\{i \in Q_s(\mathbf{c}_n) \text{ such that } \pi_{is} \leq \tilde{c}_{ns}\}) + \tilde{f}_n(\{i \in Q_s(\mathbf{c}_n) \text{ such that } \tilde{c}_{ns} < \pi_{is} \leq c_s\}) \\
&\quad = \tilde{f}_n(\{i \in Q_s(\mathbf{c}_n) \text{ such that } \pi_{is} \leq \tilde{c}_{ns}\}) \qquad (22) \\
&\quad \leq q_s. \qquad (23)
\end{aligned}$$

Expression (22) follows from Assumption 1 by the following reason. (22) does not hold, i.e., $\tilde{f}_n(\{i \in Q_s(\mathbf{c}_n) \text{ such that } \tilde{c}_{ns} < \pi_{is} \leq c_s\}) > 0$ only if $F(\{i \in I | \tilde{c}_{ns} < \pi_{is} \leq c_s\}) > 0$. This and Assumption 1 imply $F(\{i \in I | \tilde{c}_{ns} < \pi_{is} \leq c_s \text{ and } i \text{ ranks } s \text{ first}\}) \equiv F(I_n^s) > 0$, a contradiction to $F(I_n^s) = 0$. Since $\tilde{f}_n$ is realized as n iid samples from F, $\tilde{f}_n(\{i \in I | \tilde{c}_{ns} < \pi_{is} \leq c_s\}) = 0$. Expression (23) follows by our definition of DA, which can never assign more students to a school than its capacity for each of the n samples. We obtain our contradiction since $\tilde{c}_{ns}$ is not maximal at s in $\tilde{f}_n$ since expression (23) means it is possible to increase the cutoff $\tilde{c}_{ns}$ to c_s without violating the capacity constraint.

Given that we've just shown that for each $s \in S'$, $F(I_n^s) > 0$ for some n, it is possible to find an n such that $F(I_n^s) > \epsilon > 0$. Since $f_n \to F$ and so $\tilde{f}_n \to F$, there exists N such that for all $n > N$, we have $\tilde{f}_n(I_n^s) > F(I_n^s) - \epsilon > 0$. Since the number of schools is finite, such N can be taken uniformly over all $s \in S$. This completes the argument for Step (a).

Step (a) allows us to find some N such that for any $n > N$, $\tilde{f}_n(I_n^s) > 0$ for all $s' \in S'$. Let $\tilde{s}_n \in S$ and t be such that $\tilde{c}_{n\tilde{s}}^{t-1} \geq c_s$ for all $s \in S$ and $\tilde{c}_{n\tilde{s}_n}^{t} < c_{\tilde{s}_n}$. That is, $\tilde{s}_n$ is one of the first schools the cutoff of which falls strictly below $c_{\tilde{s}_n}$ under the DA algorithm in $\tilde{f}_n$, which happens in round t of the DA algorithm. Such $\tilde{s}_n$ and t exist since the choice of n guarantees $\tilde{f}_n(I_n^s) > 0$ and so $\tilde{c}_{ns} < c_s$ for all $s \in S'$.

Step (b): We next show that there exist infinitely many values of n such that the associated $\tilde{s}_n$ is in S' and $\tilde{f}_n(I_n^s) > 0$ for all $s \in S'$. It is because otherwise, by Step (a), there exists N such that for all $n > N$, we have $\tilde{s}_n \notin S'$. Since there are only finitely many schools, $\{\tilde{s}_n\}$ has a subsequence $\{\tilde{s}_m\}$ such that $\tilde{s}_m$ is the same school outside S' for all m. By definition of $\tilde{s}_n$, $\tilde{c}_{m\tilde{s}_m} \leq \tilde{c}^t_{m\tilde{s}_m} < c_{\tilde{s}_m}$ for all m and so $\tilde{c}_{\tilde{s}_m} < c_{\tilde{s}_m}$, a contradiction to $\tilde{s}_m \notin S'$. Therefore, we have our desired conclusion of Step (b).

Fix some n such that the associated $\tilde{s}_n$ is in S' and $\tilde{f}_n(l_n^s) > 0$ for all $s \in S'$. Step (b) guarantees that such n exists. Let $\tilde{A}_{n\tilde{s}_n}$ and $A_{\tilde{s}_n}$ be the sets of students assigned $\tilde{s}_n$ under $\tilde{f}_n$ and F, respectively. All students in $I_n^{\tilde{s}_n}$ are assigned $\tilde{s}_n$ in F and rejected by $\tilde{s}_n$ in $\tilde{f}_n$. Since these students rank $\tilde{s}_n$ first, there must exist a positive measure (with respect to $\tilde{f}_n$) of students outside $I_n^{\tilde{s}_n}$ who are assigned $\tilde{s}_n$ in $\tilde{f}_n$ and some other school in F; denote the set of them by $\tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$. $\tilde{f}_n(\tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}) > 0$ since otherwise, for any n such that Step (b) applies,

$$\tilde{f}_n(\tilde{A}_{n\tilde{s}_n}) \leq \tilde{f}_n(A_{\tilde{s}_n} \setminus I_n^{\tilde{s}_n}) = \tilde{f}_n(A_{\tilde{s}_n}) - \tilde{f}_n(I_n^{\tilde{s}_n}),$$

which by Step (a) converges to something strictly smaller than $F(A_{\tilde{s}_n})$ since $\tilde{f}_n(A_{\tilde{s}_n}) \to F(A_{\tilde{s}_n})$ and $\tilde{f}_n(I_n^{\tilde{s}_n}) > 0$ for all large enough n by Step (a). Note that $F(A_{\tilde{s}_n})$ is weakly smaller than $q_{\tilde{s}_n}$. This implies that for large enough n, $\tilde{f}_n(\tilde{A}_{n\tilde{s}_n}) < q_{\tilde{s}_n}$, a contradiction to $\tilde{A}_{n\tilde{s}_n}$'s being the set of students assigned $\tilde{s}_n$ at a cutoff strictly smaller than the largest possible value $K+1$. For each $i \in \tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$, let s_i be the school to which i is assigned under F.

Step (c): To complete the argument for Claim 1, we show that some $i \in \tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$ must have been rejected by s_i in some step $\tilde{t} \leq t-1$ of the DA algorithm in $\tilde{f}_n$. That is, there exists $i \in \tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$ and $\tilde{t} \leq t-1$ such that $\pi_{is_i} > \tilde{c}^{\tilde{t}}_{ns_i}$. Suppose to the contrary that for all $i \in \tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$ and $\tilde{t} \leq t-1$, we have $\pi_{is_i} \leq \tilde{c}^{\tilde{t}}_{ns_i}$. Each such student i must prefer s_i to $\tilde{s}_n$ because i is assigned $s_i \neq \tilde{s}_n$ under F though $\pi_{i\tilde{s}_n} \leq \tilde{c}_{n\tilde{s}_n} < c_{\tilde{s}_n}$, where the first inequality holds because i is assigned $\tilde{s}_n$ in $\tilde{F}_n$ while the second inequality does because $\tilde{s}_n \in S'$. This implies none of $\tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$ is rejected by s_i, applies for $\tilde{s}$, and contributes to decreasing $\tilde{c}^t_{n\tilde{s}_n}$ at least until step t and so $\tilde{c}^t_{n\tilde{s}_n} < c_{\tilde{s}_n}$ cannot be the case, a contradiction. Therefore, we have our desired conclusion of Step (c).

Claim 1 can now be established by showing that Step (c) implies there are $i \in \tilde{A}_{n\tilde{s}_n} \setminus A_{\tilde{s}_n}$ and $\tilde{t} \leq t-1$ such that $\pi_{is_i} > \tilde{c}^{\tilde{t}}_{ns_i} \geq \tilde{c}_{ns_i}$, where the last inequality is implies by the fact that in every economy, for all $s \in S$ and $t \geq 0$, we have $c_s^{t+1} \leq c_s^t$. Also, they are assigned s_i in F so that $\pi_{is_i} \leq c_{s_i}$. These imply $c_{s_i} > \tilde{c}^{\tilde{t}}_{ns_i} \geq \tilde{c}_{ns_i}$. That is, the cutoff of s_i falls below c_{si} in step $\tilde{t} \leq t-1 < t$ of the DA algorithm in $\tilde{f}_n$. This contradicts the definition of $\tilde{s}_n$ and t. Therefore $\tilde{c}_s \geq c_s$ for all $s \in S$, as desired. □

Claim 2. *By a similar argument, $\tilde{c}_s \leq c_s$ for every $s \in S$.*

Since $\tilde{c}_s \geq c_s$ and $\tilde{c}_s \leq c_s$ for all s, it must be the case that $\tilde{\mathbf{c}}_n \to \mathbf{c}$. The following claim uses this to show that $\mathbf{c}_n \to \mathbf{c}$.

Claim 3. *If $\tilde{\mathbf{c}}_n \to \mathbf{c}$ for every convergent subsequence $\{\tilde{\mathbf{c}}_n\}$ of $\{\mathbf{c}_n\}$, then $\mathbf{c}_n \to \mathbf{c}$.*

Proof of Claim 3. Since $\{\mathbf{c}_n\}$ is bounded in $[0, K+1]^{|S|}$, it has a convergent subsequence by the Bolzano-Weierstrass theorem. Suppose to the contrary that for every convergent subsequence $\{\tilde{\mathbf{c}}_n\}$, we have $\tilde{\mathbf{c}}_n \to \mathbf{c}$, but $\mathbf{c}_n \not\to \mathbf{c}$. Then there exists $\epsilon > 0$ such that for all $k > 0$, there exists $n_k > k$ such that $||\mathbf{c}_{n_k} - \mathbf{c}|| \geq \epsilon$. Then the subsequence $\{\mathbf{c}_{n_k}\}_k \subset \{\mathbf{c}_n\}$ has a convergent subsequence that does not converge to $\mathbf{c}$ (since $||\mathbf{c}_{n_k} - \mathbf{c}|| \geq \epsilon$ for all k), which contradicts the supposition that every convergent subsequence of $\{\mathbf{c}_n\}$ converges to $\mathbf{c}$. □

The last step in the proof of Lemma 1 relates this fact to stochastic convergence.

Claim 4. $\mathbf{c}_n \to \mathbf{c}$ *implies* $\hat{\mathbf{c}}_n \xrightarrow{a.s.} \mathbf{c}$

Proof of Claim 4. This proof is based on two off-the-shelf asymptotic results from mathematical statics. First, let F_n be the distribution over $I(\Theta_0, r_0, r_1)$'s generated by randomly drawing n students from F. Note that F_n is random since it involves randomly drawing n students. $F_n \xrightarrow{a.s.} F$ by the Glivenko-Cantelli theorem (Theorem 19.1 in van der Vaart (2000)). Next, since $F_n \xrightarrow{a.s.} F$ and $\mathbf{c}_n \to \mathbf{c}$, the Extended Continuous Mapping Theorem (Theorem 18.11 in van der Vaart (2000)) implies that $\hat{\mathbf{c}}_n \xrightarrow{a.s.} \mathbf{c}$, completing the proof of Lemma 1. □

A.4.2 Proof of Lemma 2

Proof. Consider any deterministic sequence of economies $\{f_n\}$ such that $f_n \in \mathcal{F}$ for all n and $f_n \to F$ in the $(\mathcal{F}, d)$ metric space. Let $p_{ns}(\theta)$ be the (finite-market, deterministic) propensity score for a particular f_n. Note that this subtly modifies the definition of $p_{ns}(\theta)$ from that in the text. The change here is that the propensity score for f_n is not a random quantity, because economy f_n is viewed as fixed.

For Lemma 2, it is enough to show deterministic convergence of this finite-market score, that is, $p_{ns}(\theta) \to \varphi_s(\theta)$ as $f_n \to F$. To see this, let F_n be the distribution over $I(\Theta_0, r_0, r_1)$'s induced by randomly drawing n students from F. Note that F_n is random and that $F_n \xrightarrow{a.s.} F$ by the Glivenko-Cantelli theorem (Theorem 19.1 in van der Vaart (2000)). $F_n \xrightarrow{a.s.} F$ and $p_{ns}(\theta) \to \varphi_s(\theta)$ allow us to apply the Extended Continuous Mapping Theorem (Theorem 18.11 in van der Vaart (2000)) to obtain $\tilde{p}_{ns}(\theta) \xrightarrow{a.s.} \varphi_s(\theta)$.

We prove convergence of $p_{ns}(\theta) \to \varphi_s(\theta)$ as follows. Let $\tilde{c}_{ns}$ and $\tilde{c}_{ns'}$ be the random cutoffs at s and s', respectively, in f_n, and

$\tau_{\theta s} \equiv c_s - \rho_{\theta s}$,
$\tau_{\theta s\_} \equiv \max_{s' \succ_\theta s}\{c_{s'} - \rho_{\theta s'}\}$,
$\tilde{\tau}_{n\theta s} \equiv \tilde{c}_{ns} - \rho_{\theta s}$, and
$\tilde{\tau}_{n\theta s\_} \equiv \max_{s' \succ_\theta s}\{\tilde{c}_{ns'} - \rho_{\theta s'}\}$.

We can express $\varphi_s(\theta)$ and $p_{ns}(\theta)$ as follows.

$\varphi_s(\theta) = \max\{0, \tau_{\theta s} - \tau_{\theta s\_}\}$
$p_{ns}(\theta) = P_n(\tilde{\tau}_{n\theta s} \geq R > \tilde{\tau}_{n\theta s\_})$

where P_n is the probability induced by randomly drawing lottery numbers given f_n, and R is any type θ student's random lottery number distributed according to $U[0,1]$. By Lemma 1, with probability 1, for all $\epsilon_1 > 0$, there exists N_1 such that for all $n > N_1$,

$$|\tilde{c}_{ns'} - c_{s'}| < \epsilon_1 \text{ for all } s',$$

which implies that with probability 1,

$$\begin{aligned}
&|\tilde{\tau}_{n\theta s_-} - \tau_{\theta s_-}| \\
=&|\{\tilde{c}_{ns_1} - \rho_{\theta s_1}\} - \{c_{s_2} - \rho_{\theta s_2}\}| \\
<&\begin{cases} |\{\tilde{c}_{ns_1} - \rho_{\theta s_1}\} - (\{\tilde{c}_{ns_1} - \rho_{\theta s_2}\} + \epsilon_1)| \text{ if } c_{s_2} - \rho_{\theta s_2} \geq \tilde{c}_{ns_1} - \rho_{\theta s_1} \\ |\{\tilde{c}_{ns_1} - \rho_{\theta s_1}\} - (\{\tilde{c}_{ns_1} - \rho_{\theta s_2}\} - \epsilon_1)| \text{ if } c_{s_2} - \rho_{\theta s_2} < \tilde{c}_{ns_1} - \rho_{\theta s_1} \end{cases} \\
=&\ \epsilon_1
\end{aligned}$$

where in the first equality, $s_1 \equiv \arg\max_{s' \succ_\theta s}\{\tilde{c}_{ns'} - \rho_{\theta s'}\}$ and $s_2 \equiv \arg\max\{c_{s'} - \rho_{\theta s'}\}$. The inequality is by $|\tilde{c}_{ns'} - c_{s'}| < \epsilon_1$ for all s'. For all $\epsilon > 0$, the above argument with setting $\epsilon_1 < \epsilon/2$ implies that there exists N such that for all $n > N$,

$$\begin{aligned}
&p_{ns}(\theta) \\
&= P_n(\tilde{\tau}_{n\theta s} \geq R > \tilde{\tau}_{n\theta s_-}\}) \\
&\in (\max\{0, \tau_{\theta s} - \tau_{\theta s_-} - \epsilon, \max\{0, \tau_{\theta s} - \tau_{\theta s_-} + \epsilon) \\
&\in (\varphi_s(\theta) - \epsilon, \varphi_s(\theta) + \epsilon),
\end{aligned}$$

where the second-to-last inclusion is because with probability 1, there exists N such that for all $n > N$ such that $|\tilde{\tau}_{n\theta s} - \tau_{\theta s}|, |\tilde{\tau}_{n\theta s_-} - \tau_{\theta s_-}| < \epsilon_1$ and $R \sim U[0,1]$. This means $p_{ns}(\theta) \to \varphi_s(\theta)$, completing the proof of Lemma 2. □

A.5 First Choice and Qualification Instruments: Details

Let D_i^f be the first choice instrument defined in section 4.5 and let $\tilde{s}_i$ be i's first choice school. The first choice risk set is $Q(\theta_i) \equiv (\tilde{s}_i, \rho_{i\tilde{s}})$.

Proposition 3. *In any continuum economy, D_i^f is independent of θ_i conditional on $R(\theta_i)$.*

Proof. In general,

$$\begin{aligned}
&\Pr(D_i^f = 1|\theta_i = \theta) \\
&= \Pr(\pi_{i\tilde{s}_i} \leq c_{\tilde{s}_i}|\theta_i = \theta) \\
&= \Pr(\rho_{i\tilde{s}_i} + r_i \leq c_{\tilde{s}_i}|\theta_i = \theta) \\
&= \Pr(r_i \leq c_{\tilde{s}_i} - \rho_{i\tilde{s}_i}|\theta_i = \theta) \\
&= c_{\tilde{s}_i} - \rho_{i\tilde{s}_i},
\end{aligned}$$

which depends on θ_i only through $R(\theta_i)$ because cutoffs are fixed in the continuum.. □

Let D_i^q and $R(\theta_i)$ be the qualification instrument and the associated risk set defined in section 4.5. The latter is given by the list of schools i ranks and his priority status at each, that is, $R(\theta_i) \equiv (S_i, (\rho_{is})_{s \in S_i})$ where S_i is the set of charter schools i ranks.

Proposition 4. *In any continuum economy, D_i^q is independent of θ_i conditional on $R(\theta_i)$.*

Proof. In general, we have

$\Pr(D_i^q = 1|\theta_i = \theta)$
$= \Pr(\pi_{is} \leq c_s \text{ for some } s \in S_i|\theta_i = \theta)$
$= \Pr(\rho_{is} + r_i \leq c_s \text{ for some } s \in S_i|\theta_i = \theta)$
$= \Pr(r_i \leq c_s - \rho_{is} \text{ for some } s \in S_i|\theta_i = \theta)$
$= \Pr(r_i \leq \max_{s \in S_i}(c_s - \rho_{is})|\theta_i = \theta)$
$= \max_{s \in S_i}(c_s - \rho_{is})$,

which depends on θ_i only through $R(\theta_i)$ because cutoffs are fixed in the continuum. □

A.6 Extension to a General Lottery Structure

Washington DC, New Orleans, and Amsterdam use DA with multiple lottery numbers, one for each school (see, for example, de Haan et al. (2015)). Washington, DC uses a version of DA that uses a mixture of shared and individual school lotteries. This section derives the DA propensity score for a mechanism with any sort of multiple tie-breaking.

Let a random variable R_{is} denote student i's lottery number at school s. Assume that each R_{is} is drawn from $U[0,1]$, independently with schools. We consider a general lottery structure where $R_{is} \neq R_{is'}$ for some (not necessarily all) $s, s' \in S$ and $i \in I$.

Recall $B_{\theta s}$ is defined as $\{s' \in S \mid s' \succ_\theta s\}$. Partition $B_{\theta s}$ into $\bar{m}$ disjoint sets $B_{\theta s}^1, ..., B_{\theta s}^{\bar{m}}$, so that s' and s'' use the same lottery if and only if $s', s'' \in B_{\theta s}^m$ for some m. Note that this partition is specific to type θ. With single-school lotteries, $\bar{m}$ simplifies to $|B_{\theta s}|$, the number of schools type θ ranks ahead of s.

The *most informative disqualification*, $MID_{\theta s}^m$, is defined for each m as

$$MID_{\theta s}^m \equiv \begin{cases} 0 & \text{if } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ for all } \tilde{s} \in B_{\theta s}^m, \\ 1 & \text{if } \rho_{\theta\tilde{s}} < \rho_{\tilde{s}} \text{ for some } \tilde{s} \in B_{\theta s}^m, \\ \max\{\tau_{\tilde{s}} \mid \tilde{s} \in B_{\theta s}^m \text{ and } \rho_{\theta\tilde{s}} = \rho_{\tilde{s}}\} & \text{if } \rho_{\theta\tilde{s}} = \rho_{\tilde{s}} \text{ for } \tilde{s} \in B_{\theta s}^m \text{ and } \rho_{\theta\tilde{s}} > \rho_{\tilde{s}} \text{ otherwise.} \end{cases}$$

Let m^* be the value of m for schools in the partition that use the same lottery as s. Denote the associated MID by $MID_{\theta s}^*$. We define $MID_{\theta s}^* = 0$ when the lottery at s is unique and there is no m^*. The following result extends Theorem 1 to a general lottery structure. The proof is omitted.

Theorem 1 (Generalization). *For all s and θ in any continuum economy, we have:*

$$\Pr[D_i(s) = 1|\theta_i = \theta] = \varphi_s(\theta) \equiv \begin{cases} 0 & \text{if } \theta \in \Theta_s^n, \\ \Pi_{m=1}^{\bar{m}}(1 - MID_{\theta s}^m) & \text{if } \theta \in \Theta_s^a, \\ \Pi_{m=1}^{\bar{m}}(1 - MID_{\theta s}^m) \times \max\left\{0, \dfrac{\tau_s - MID_{\theta s}^*}{1 - MID_{\theta s}^*}\right\} & \text{if } \theta \in \Theta_s^c. \end{cases}$$

where we set $\varphi_s(\theta) = 0$ when $MID_{\theta s}^* = 1$ and $\theta \in \Theta_s^c$.

Note that in the single tie breaker case, the expression for $\varphi_s(\theta)$ reduces to that in Theorem 1 since $\bar{m} = 1$ in that case.

A.7 The Boston (Immediate Acceptance) Mechanism

Studies by Hastings-Kane-Staiger (2009), Hastings-Neilson-Zimmerman (2012), and Deming-Hastings-Kane-Staiger (2013)), among others, use data generated from versions of the Boston mechanism. Given strict preferences of students and schools, the Boston mechanism is defined as follows:

- Step 1: Each student applies to her most preferred acceptable school (if any). Each school accepts its most-preferred students up to its capacity and rejects every other student.

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

- Step t: Each student who has not been accepted by any school applies to her most preferred acceptable school that has not rejected her (if any). Each school accepts its most-preferred students up to its remaining capacity and rejects every other student.

This algorithm terminates at the first step in which no student applies to a school. Boston assignments differ DA in that any offer at any step is fixed; students receiving offers cannot be displaced later.

This important difference notwithstanding, the Boston mechanism can be represented as a special case of DA by redefining priorities as follows:

Proposition 5. *(Ergin and Sönmez (2006)) The Boston mechanism applied to $(\succ_i)_i$ and $(\succ_s)_s$ produces the same assignment as DA applied to $(\succ_i)_i$ and $(\succ_s^*)_s$ where $\succ_s^*$ is defined as follows:*

1. *For $k = 1, 2...$, $\{$students who rank s k-th$\}$ $\succ_s^*$ $\{$students who rank s $k+1$-th$\}$*
2. *Within each category, $\succ_s^*$ ranks the students in the same order as original $\succ_s$.*

This equivalence allows us to construct a Boston propensity score by redefining priorities so that priority groups at a given school consists of applicants who (i) share the same original priority status at the school and (ii) give the same rank to the school.

References

Abadie, Alberto, "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 2002, *97(457)*, 284–292.

_ , "Semiparametric instrumental variables estimation of treatment response models," *Journal of Econometrics*, 2003, *113(2)*, 231–263.

_ **and Guido Imbens**, "Matching on the Estimated Propensity Score," 2012. Working Paper.

_ **, Susan Athey, Guido Imbens, and Jeffrey M. Woolridge**, "Finite Population Causal Standard Errors," 2014. NBER Working Paper 20325.

Abdulkadiroglu, Atila and Tayfun Sonmez, "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems," *Econometrica*, 1998, *66(3)*, 689–701.

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

_ **, Josh Angrist, and Parag Pathak**, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 2014, *82(1)*, 137–196.

_ **, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak**, "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots," *Quarterly Journal of Economics*, 2011, *126(2)*, 699–748.

_ **, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez**, "Changing the Boston School Choice Mechanism," 2006. NBER Working paper, 11965.

_ , _ , **and** _ , "Designing the New Orleans School Match: TTC vs. DA," Working Paper, MIT 2014.

_ **, Weiwei Hu, and Parag Pathak**, "Small High Schools and Student Achievement: Lottery-Based Evidence from New York City," 2013. NBER Working paper, 19576.

_ **, Yeon-Koo Che, and Yosuke Yasuda**, "Expanding "Choice" in School Choice," *American Economic Journal: Microeconomics*, 2015, *7(1)*, 1–42.

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

Angrist, Joshua D. and Jinyong Hahn, "When to Control for Covariates? Panel Asymptotics for Estimates of Treatment Effects," *The Review of Economics and Statistics*, 2004, *86(1)*, 58–72.

_ **and Jörn-Steffen Pischke**, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

_ , **Guido W. Imbens, and Donald B. Rubin**, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 1996, *91(434)*, 444–455.

_ , **Parag A. Pathak, and Christopher R. Walters**, "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, 2013, *5(4)*, 1–27.

_ , **Sarah Cohodes, Susan Dynarski, Parag Pathak, and Christopher Walters**, "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics*, 2015, *Forthcoming.*

Azevedo, Eduardo and Jacob Leshno, "A Supply and Demand Framework for Two-Sided Matching Markets," 2014. Working Paper.

Balinski, Michel and Tayfun Sönmez, "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 1999, *84*, 73–94.

Behaghel, Luc, Bruno Crépon, and Marc Gurgand, "Robustness of the Encouragement Design in a Two-Treatment Randomized Control Trial," 2013. IZA Discussion Paper No. 7447.

Blackwell, Matthew, "Identification and Estimation of Joint Treatment Effects with Instrumental Variables," 2015. Working paper, Harvard University.

Bloom, Howard and Rebecca Unterman, "Can Small High Schools of Choice Improve Educational Prospects for Disadvantaged Students?," *Journal of Policy Analysis and Management*, 2014, *33(2).*

Budish, Eric, "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes," *Journal of Political Economy*, 2011, *119(6)*, 1061–1103.

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

Che, Yeon-Koo and Fuhito Kojima, "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms," *Econometrica*, 2010, *78(5)*, 1625–1672.

Connors, Susan, Erika Moldow, Amelia Challender, and Bonnie Walters, "Innovation Schools in DPS: Year Three of an Evaluation Study," *University of Colorado Denver: The Evaluation Center, School of Education and Human Development*, 2013.

Cullen, Jullie Berry, Brian A. Jacob, and Steven Levitt, "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, 2006, *74(5)*, 1191–1230.

de Haan, Monique, Pieter A. Gautier, Hessel Oosterbeek, and Bas van der Klaauw, "The Performance of School Assignment Mechanisms in Practice," 2015. Working Paper.

Deming, David, "Better Schools, Less Crime?," *Quarterly Journal of Economics*, 2011, *126(4)*, 2063–2115.

_ , **Justine Hastings, Thomas Kane, and Douglas Staiger**, "School Choice, School Quality and Postsecondary Attainment," *American Economic Review*, 2014, *104(3)*, 991–1013.

Dobbie, Will and Roland G. Fryer, "Exam High Schools and Academic Achievement: Evidence from New York City," *American Economic Journal: Applied Economics*, 2014, *6(3)*, 58–75.

Dobbie, William and Roland Fryer, "Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 2011, *3(3)*, 158–187.

Ergin, Haluk and Tayfun Sönmez, "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 2006, *90*, 215–237.

Gale, David and Lloyd S. Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69*, 9–15.

Hahn, Jinyong, "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, 1998, *66(2)*, 315–331.

Hastings, Justine, Thomas J. Kane, and Douglas O. Staiger, "Heterogenous Preferences and the Efficacy of Public School Choice," 2009. Working paper, Yale University.

Hirano, Keisuke, Guido Imbens, and Geert Ridder, "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, 2003, *71(4)*, 1161–1189.

Hoxby, Caroline M., Sonali Murarka, and Jenny Kang, "How New York City's Charter Schools Affect Achievement," 2009. Working Paper.

Imbens, Guido, "The Role of the Propensity Score in Estimating Dose-response Functions," *Biometrica*, 2000, *87(3)*, 706–710.

_ **and Joshua D. Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, pp. 467–475.

Jackson, Kirabo, "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago," *Economic Journal*, 2010, *120(549)*, 1399–1429.

Kesten, Onur and Utku Ünver, "A Theory of School Choice Lotteries," *Theoretical Economics*, 2015, *10(2)*, 543–595.

Lehmann, Erich L. and Joseph P. Romano, *Testing Statistical Hypotheses*, Springer, 2005.

Lucas, Adrienne and Isaac Mbiti, "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya," *American Economic Journal: Applied Economics*, 2014, *6(3)*, 234–63.

Pathak, Parag A. and Jay Sethuraman, "Lotteries in Student Assignment: An Equivalence Result," 2010. NBER Working paper 16140.

_ **and Tayfun Sönmez**, "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 2008, *98(4)*, 1636–1652.

Pop-Eleches, Cristian and Miguel Urquiola, "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 2013, *103(4)*, 1289–1324.

Rosenbaum, Paul R. and Donald B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrica*, 1983, pp. 41–55.

Roth, Alvin E. and Marilda A. O. Sotomayor, *Two-sided Matching: a Study in Game-theoretic Modeling and Analysis*, Cambridge University Press: Econometric Society monographs, 1990.

Shao, Jun and Dongsheng Tu, *The Jackknife and Bootstrap*, Springer, 1995.

Shapley, Lloyd and Herbert Scarf, "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1974, *1*, 23–28.

Svensson, Lars-Gunnar, "Strategy-proof Allocation of Indivisible Goods," *Social Choice and Welfare*, 1999, *16*.

van der Vaart, A. W., *Asymptotic Statistics*, Cambridge University Press, 2000.

Yang, Shu, Guido Imbens, Zhanglin Cui, Douglas Faries, and Zibgniew Kadziola, "Propensity Score Matching and Subclassification in Observational Studies with Multi-level Treatments," 2014. Unpublished manuscript.