# Expanding "Choice" in School Choice

Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda*

February 10, 2011

Abstract: Gale-Shapley's deferred acceptance (henceforth DA) mechanism has emerged as a prominent candidate for placing students to public schools. While the DA has desirable fairness and incentive properties, it limits the applicants' abilities to communicate their preference intensities, which results in ex-ante inefficiency when ties at school preferences are broken randomly. We propose a variant of deferred acceptance mechanism which allows students to influence how they are treated in ties. It inherits much of the desirable properties of the DA but performs better in ex ante efficiency.

Keywords: Gale-Shapley's deferred acceptance algorithm, choice-augmented deferred acceptance, tie breaking, ex ante Pareto efficiency.

*Abdulkadiroğlu: Department of Economics, Duke University, Durham, NC 27708 (email: atila.abdulkadiroglu@duke.edu); Che: Department of Economics, Columbia University, New York City, NY 10027 and YERI, Yonsei University (email: yeonkooche@columbia.edu); Yasuda: National Graduate Institute for Policy Studies, Tokyo 106-8677, Japan (email: yyasuda@grips.ac.jp). We are grateful to Francis Bloch, Olivier Compte, Haluk Ergin, Matt Jackson, Navin Kartik, Fuhito Kojima, David Levine, Paul Milgrom, Andrew Postlewaite, Al Roth, Ilya Segal, Ron Siegel, and other seminar participants at Brown University, Caltech, Duke University, GREQAM, GRIPS, Harvard Market Design Workshop, Hong Kong University, Nihon University, Northwestern University, Paris School of Economics, Princeton University (IAS), Stanford University, Toyama University, Tsukuba University, Waseda University, Washington University, Yonsei University, and Universities of Pompeu Fabra, and Rochester.

# 1 Introduction

The promise of school choice programs is to expand students' access to public schools beyond their residential boundaries.[1] In reality, this promise is tampered by the fact that public schools face physical capacity constraints. When too many students demand a seat at a particular school, admissions to the school is regulated by priorities, which may be determined by the location of the school, students' home addresses and other criteria. Priorities may serve broader public policy purpose, but they are typically very coarse. For instance, Boston Public Schools (BPS) prioritize applicants based only on sibling attendance and "walk zone;" entailing many students to fall into the same priority class. Hence, there remains a great deal of freedom in placing students according to their preferences. The fundamental question is how to accommodate the preferences when they conflict: Namely, who should get in and who should be turned away when there is excess demand for a school?

Allocating scarce resources (in the presence of conflicting preferences) is after all the fundamental role of markets. Competitive markets allocate goods efficiently based on individuals' preference intensities. If a good is overdemanded, for instance, its price is bid up so that those who are willing to pay more for the good than others end up with it. Unfortunately, monetary exchange is not a viable option in assigning seats at public schools, since it would violate the principle of free public education. But the "cardinal" efficiency of school assignment, that seats at a popular school must go to those who would lose relatively more by being assigned the next best school, remains an important issue.[2] Can the cardinal efficiency be achieved without monetary exchange in the market of school choice? We suggest that this is indeed possible, and can be done by applying the lesson from competitive markets. We show that school choice can be designed to harness the pricing function of the market and generate a more efficient outcome in the cardinal sense, and this can be done within the framework of the most popular choice mechanism, namely Gale-Shapley's student-proposing deferred acceptance (henceforth, DA) algorithm, without sacrificing much of its beneficial properties.

Since proposed by Abdulkadiroğlu and Sönmez (2003), DA has emerged as one of the most prominent candidates for school choice design. In 2003, New York City Department of Educa-

[1]Government policies promoting school choice take various forms, including interdistrict and intradistrict public school choice as well as open enrollment, tax credits and deductions, education savings accounts, publicly funded vouchers and scholarships, private voucher programs, contracting with private schools, home schooling, magnet schools, charter schools and dual enrollment. See an interactive map at http://www.heritage.org/research/Education/SchoolChoice/SchoolChoice.cfm for a comprehensive list of choice plans throughout the US.

[2]We use the term "cardinal" in order to distinguish the efficiency we focus on from the ordinal efficiency. This distinction is made clearer by the example introduced shortly and by subsequent sections.

tion adopted the DA. In 2005, the BPS also adopted DA in place of the existing priority rule known as the "Boston" mechanism. Beyond school choice, DA has a celebrated history, having been successfully applied for matching doctors to hospital for their internships and residencies (see Roth, 2008). The algorithm works as follows. Once students submit their ordinal rankings of schools and school priorities are determined, the algorithm iterates the following procedure in successive rounds: Each student applies to her most preferred school that has not rejected her yet; each school tentatively admits up to its capacity in its priority order from the list of new applicants and those it has kept from the previous round and it rejects the remaining students. The tentative assignment becomes final when no student is rejected.

The DA has several desirable properties. It ensures fairness by eliminating justified envy, that is, no student ever loses a seat at a desired school to somebody with a lower priority at that school (Gale and Shapley, 1962; Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003). The assignment is also constrained efficient; namely, no other fair assignment can make all students better off, some strictly.[3] Furthermore, DA is strategy-proof, meaning that the students have a dominant strategy of revealing their preference rankings of schools truthfully (Dubins and Freedman, 1981; Roth, 1982).[4]

Nevertheless, the DA does not respond to applicants' cardinal preferences. When two applicants tie in priority, the DA randomly determines who will be admitted and who will be rejected. That is, when priorities do not determine the allocation, DA completely ignores the underlying preference intensities of students. Luck of "draw" instead determines an applicant's fate. This is in contrast with the competitive markets where the agents can express their preference intensities via their willingness to pay. This apparent "irony" was not lost in the run up to the BPS' adoption of the DA when parents observed:

> ... if I understand the impact of Gale Shapley, and I've tried to study it and I've met with BPS staff... I understood that in fact the random number ... [has] preference over your choices... (Recording from the BPS Public Hearing, 6-8-05).
>
> I'm troubled that you're considering a system that takes away the little power that parents have to prioritize... what you call this strategizing as if strategizing is a dirty word... (Recording from the BPS Public Hearing, 5-11-04).

This lack of responsiveness to cardinal preferences, or the inability by the parents to influence

[3]The constrained efficiency result with strict priorities follow from Gale and Shapley (1962) and takes a stronger claim: The DA assignment Pareto dominates all other fair assignments. With coarse priorities, the constrained efficiency result follows from Abdulkadiroğlu et al. (2009) and Erdil and Ergin (2008).

[4]Aside from the ease with which parents make their choice, strategy-proofness "levels the playing field," by putting those strategically unsophisticated at no disadvantage relative to those who are more sophisticated.

how they are treated at a tie, entails real welfare loss. To illustrate, suppose there are three students, $\{1, 2, 3\}$, and three schools, $\{a, b, c\}$, each with one seat. Schools have no intrinsic priorities over students, and students' preferences are represented by the following von-Neumann Morgenstern (henceforth, vNM) utility values, where $v_j^i$ is student $i$'s vNM utility value for school $j$:

| | $v_j^1$ | $v_j^2$ | $v_j^3$ |
|---|---|---|---|
| $j = a$ | 4 | 4 | 3 |
| $j = b$ | 1 | 1 | 2 |
| $j = c$ | 0 | 0 | 0 |

Since schools have no priorities, the DA algorithm generates the schools' rankings by a uniform lottery; that is, ties among students are broken randomly. By strategy-proofness, all three students submit truthful rankings of the schools. Consequently, the students are assigned each school with probability 1/3. In other words, the DA mechanism reduces to a pure lottery assignment. The students obtain expected utility of $EU_1^{DA} = EU_2^{DA} = EU_3^{DA} = \frac{5}{3}$. This assignment makes no distinction between students 1, 2 and student 3, despite the fact that the first two would suffer more by being assigned the next best alternative $b$ than student 3.

Not surprisingly, reallocating the assignment probabilities can make all agents strictly better off. Suppose instead student 3 is assigned school $b$ for sure, and students 1 and 2 are assigned between $a$ and $c$ with equal probability 1/2. Each student obtains the expected utility of $EU_1^B = EU_2^B = EU_3^B = 2$ strictly higher than $\frac{5}{3}$ they enjoyed in the DA. Moreover, this new assignment can be implemented in the incentive compatible way. Suppose for instance that the students are offered the two different lotteries; sure assignment at school $b$ versus a uniform lottery between $a$ and $c$. Intuitively, this mechanism offers a 1/2 chance of getting assigned the best school $a$, for the price of 1/2 chance of getting assigned the worst school $c$. The first two students (who will suffer a lot by being placed to the second best school instead of the best) will pay the price, and the third student (who will not suffer as much) will not. Hence, this mechanism implements the desirable assignment.

We propose a way to modify the DA algorithm to harness the "pricing" feature of this mechanism. The idea is to allow the students to signal their cardinal preferences by sending an additional message that is used to break ties in DA. In what we call "Choice-Augmented Deferred Acceptance" (CADA) algorithm, each student submits an ordinal list of schools just as before, but she also submits the name of a "target" school. A school then elevates the priority standing of those who targeted that school and favor them in breaking ties among those with the same priority at the school. The iterative procedure of DA is then implemented with the rankings generated in this way. The CADA inherits the desirable properties of the standard

DA: It is fair in the sense of eliminating justified envy, strategy-proof with respect to the ordinal preference lists, and constrained efficient. Clearly, targeting involves strategic behavior, but its importance is limited by the priorities. If schools priorities are strict, then there are no ties, so CADA coincides with the standard DA.

If the priorities are coarse, then targeting allows individuals to signal their preference intensities, and in the process serves to "price" the schools based on their demands. Intuitively, if a school is targeted by more students, one finds it more difficult to raise the odds of assignment at that school via targeting that school; the price of that school has effectively risen. This feature of CADA allows competitive markets to operate for a set of popular schools, attaining ex ante efficiency within these schools. For this reason, as we show formally in the sequel, for a large economy (both in the size of student body and school capacities), the CADA performs better than the DA with standard random tie-breaking rules in ex ante welfare. For instance, in the above example, the unique Nash equilibrium of the CADA has students 1 and 2 will target $a$, and student 3 will target $b$, so the desirable outcome is implemented.[5]

The issue of cardinal welfare, and the ex ante efficiency that captures the cardinal welfare, has not received much attention in the debate of school choice design. The existing debate has largely focused on ex post efficiency or ordinal efficiency as an welfare concept.[6] The cardinal welfare issue would not matter much if either students' preferences are diverse or if the schools' priorities are strict. In the former case, the preferences do not conflict, so they can be easily accommodated. In the latter case, even if the preferences conflict, school priorities pin down the assignment, so there is no scope for assignment to respond to cardinal welfare.

If neither is true, however, the cardinal welfare issue becomes important. Suppose for an extreme case that the agents have the same ordinal preferences and schools have no priorities. In that case, as seen in the example, it matters how the agents are assigned based on their relative preference intensities of the schools. By contrast, ex post efficiency loses its bite; as can be seen in the example, *all* assignments are ex post efficient, thus indistinguishable on this ground. In particular, the celebrated DA reduces to a pure lottery assignment. The mechanism such as CADA can do strictly better in such a case. Although this extreme scenario is special, it captures the salient features of the reality. As noted, priorities tend to be coarse

[5]All students will submit their rankings truthfully, so $a - b - c$ in this order. School $a$ will then rank students 1 and 2 ahead of 3 (but randomly between the first two), and school $b$ will rank student 3 ahead of 1 and 2 (again randomly between these two). In the first round of CADA, all students apply to $a$ and it will choose between 1 and 2, and the two rejected students, including student 3, will then apply to school $b$ in the second round, and school $b$ will admit student 3. Hence, student 3 is assigned $b$ for sure, and 1 and 2 are assigned between $a$ and $c$ with equal probability, thus implementing the superior assignment.

[6]See for example Erdil and Ergin (2008) and Kesten (2010).

and parents tend to value the similar qualities about schools (e.g., safety, academic excellence, etc.) in reality, leading them to have similar preferences. Indeed, the BPS data exhibits strong correlation across students' ordinal preferences over schools. In 2007-2008, only 8 out of 26 schools (at grade level 9) were overdemanded whereas an average of 22.21 (std 0.62) schools should have been overdemanded if students' preferences had been uncorrelated.[7] As we point out in this paper, in such an environment, CADA performs particularly well relative to the DA with standard random tie-breaking rules.

The rest of the paper is organized as follows. Section 2 defines CADA more precisely and shows that it inherits the desired properties of standard DA. Section 3 introduces the formal model and welfare criterion. Section 4 provides welfare comparison across the three alternative procedures. Section 5 presents simulation to quantify the welfare benefits of CADA. Section 6 then considers the implication of enriching the message used in the CADA and the robustness of our results to some students not behaving in a strategically sophisticated way. Section 7 concludes. All proofs that do not appear in this paper are available in the Supplementary Notes ("not for publication").

# 2 Choice-Augmented DA Algorithm: Definition

Here we describe the DA more precisely when schools priorities are not strict. DA requires the rankings to be strict on both sides, which means that ties must be broken to *generate* strict preferences on the school side. There are two common methods of breaking ties. *Single tie-breaking* randomly assigns every student a single lottery number to break ties at every school, whereas *multiple tie-breaking* randomly assigns a distinct lottery number to each student at every school. Clearly, a DA algorithm is well defined with respect to the strict priority list generated by either method. We refer the DA algorithms using single and multiple tie-breaking by *DA-STB* and *DA-MTB*, respectively.

We propose an alternative way to break a tie, one that allows students to influence its outcome based on their communication. The associated DA algorithm, which we refer to as *Choice-Augmented Deferred Acceptance* (henceforth, CADA), proceeds as follows:

- **Step 1:** All students submit ordinal preferences, plus an "auxiliary message," naming a

[7] This comparison is based on submitted preferences under the DA introduced in 2005. Since the DA is strategy-proof and BPS paid significant attention in communicating that feature of the DA to the public, we assume that those submitted preferences are a good approximation of the underlying true preferences. For the counter-factual, we generated 100 different preference profiles by drawing a school as first choice for each student uniformly randomly from the set of schools and computed the number of overdemanded schools given school capacities.

"target" school. If a student names a school for a target, she is said to have "targeted" the school.

• **Step 2:** The schools' strict priorities over students are generated based on their *intrinsic priorities* and the students' auxiliary messages as follows. First, each student is independently randomly assigned two lottery numbers. Call one *target lottery number* and the other *regular lottery number*. Each school's *strict priority list* is then generated as follows: (i) First consider the students in the school's highest priority group. Within that group, rank at the top those who name the school as their target. List them in the order of their target lottery numbers, and list below them the rest (who didn't target that school) according to their regular lottery numbers. (ii) Move to the next highest priority group, list them below in the same fashion, and repeat this process until all students are ranked in a strict order.

• **Step 3:** The students are then assigned to schools via the DA algorithm, using *each student's ordinal preferences* from Step 1 and *each school's strict priority list* compiled in Step 2.

To illustrate Step 2, suppose there are five students $N = \{1, 2, 3, 4, 5\}$ and two schools $S = \{a, b\}$, neither of which has intrinsic priority ordering over the students. Suppose students $1, 3$ and 4 targeted $a$ and 2 and 5 targeted $b$, and that students are ordered according to their target and regular lottery numbers as follows:

$$\mathbf{T}(N) : 3 - 5 - 2 - 1 - 4; \quad \mathbf{R}(N) : 3 - 4 - 1 - 2 - 5.$$

Then the priority list for school $A$ first reorders students $\{1, 3, 4\}$, who targeted that school, based on $\mathbf{T}(N)$, to $3 - 1 - 4$, and reorders the rest, $\{2, 5\}$, based on $\mathbf{R}(N)$ to $2 - 5$, which produces a complete list for $a$:

$$\mathbf{P}_a(N) = 3 - 1 - 4 - 2 - 5.$$

Similarly, the priority list for $b$ is:

$$\mathbf{P}_b(N) = 5 - 2 - 3 - 4 - 1.$$

The process of compiling the priority lists resembles the STB in that the same lottery is used by different schools, but only within each group. Unlike STB, though, different lotteries are used across different groups. This ensures that a student who has a bad draw at her target school gets a "new lease on life" with another independent draw for the other schools.[8]

[8]Although the primary reason for having two separate random lists is to ensure a certain welfare property (see footnote 10), it has the additional benefit of allaying a concern raised in the wake of the NYC school choice redesign. One criticism against STB was that if a student has a bad draw, then she will have a bad draw with *every* school she applies to. Having two separate random lists mitigates this problem.

# 3 Model and Basic Analysis

## 3.1 Primitives

There are $n \geq 2$ schools, $S = \{s_1, s_2, ..., s_n\}$ with the index set $A := \{1, 2, ..., n\}$, each with a unit mass of seats to fill. There are mass $n$ of students who are indexed by vNM values $\mathbf{v} = (v_1, ..., v_n) \in \mathcal{V} := [0, 1]^n$ they attach to the $n$ schools. The set of student types, $\mathcal{V}$, is equipped with a measure $\mu$. We assume that $\mu$ admits strictly positive density in the interior of $\mathcal{V}$. The assumptions that the aggregate measure of students equal aggregate capacities of schools and that all students find every school acceptable are made for convenience and will not affect our main results (see Subsection 6.5).

The students' vNM values induce their ordinal preferences. Let $\pi := (\pi_1, ..., \pi_n) : \mathcal{V} \to S^n$ be such that $\pi_a(\mathbf{v}) \neq \pi_b(\mathbf{v})$ if $a \neq b$ and that $v_{\pi_a(\mathbf{v})} > v_{\pi_b(\mathbf{v})}$ implies $a < b$. In other words, $\pi(\mathbf{v})$ lists the schools in the descending order of the preferences for a student with $\mathbf{v}$, with $\pi_a(\mathbf{v})$ denoting her $a$-th preferred school. Let $\Pi$ denote the set of all ordered lists of $S$. Then, for each $\tau \in \Pi$,

$$m_\tau := \mu(\{\mathbf{v} | \pi(\mathbf{v}) = \tau\})$$

represents the measure of students whose ordinal preferences are $\tau$. By the full support assumption, $m_\tau > 0$ for each $\tau \in \Pi$. Finally, let $\mathfrak{m} := \{m_\tau\}_{\tau \in \Pi}$ be a profile of measures of all ordinal types. Let $\mathfrak{M} := \{\{m_\tau\}_{\tau \in \Pi} | \sum_{\tau \in \Pi} m_\tau = n\}$ be the set of all possible measure profiles. We say a property holds *generically* if it holds for a subset of $\mathfrak{m}$'s that has the same Lebesque measure as $\mathfrak{M}$.

An *assignment*, denoted by $\mathbf{x}$, is a probability distribution over $S$, and this is an element of a simplex, $\Delta := \{(x_1, ..., x_n) \in \mathbb{R}^n_+ | \sum_{a \in A} x_a = 1\}$. We are primarily interested in how a procedure determines the assignment for each student *ex ante*, prior to conducting the lottery. To this end, we define an *allocation* to be a measurable function $\phi := (\phi_1, ..., \phi_n) : \mathcal{V} \mapsto \Delta$ such that $\int \phi_a(\mathbf{v}) d\mu(\mathbf{v}) = 1$ for each $a \in A$, with the interpretation that student $\mathbf{v}$ is assigned by mapping $\phi = (\phi_1, ..., \phi_n)$ to school $s_a$ with probability $\phi_a(\mathbf{v})$. Let $\mathcal{X}$ denote the set of all allocations.

## 3.2 Welfare Standards

To begin, we say allocation $\tilde{\phi} \in \mathcal{X}$ *weakly Pareto dominates* allocation $\phi \in \mathcal{X}$ if, for almost every $\mathbf{v}$,

$$\sum_{a \in A} v_a \tilde{\phi}_a(\mathbf{v}) \geq \sum_{a \in A} v_a \phi_a(\mathbf{v}), \tag{1}$$

and that $\tilde{\phi}$ *Pareto dominates* $\phi$ if the former weakly dominates the latter and if the inequality of (1) is strict for a positive measure of $\mathbf{v}$'s. We also say $\tilde{\phi} \in \mathcal{X}$ *ordinally dominates* $\phi \in \mathcal{X}$ if the former has higher chance of assigning each student to her more preferred school than the latter in the sense of first-order stochastic dominance: for a.e. $\mathbf{v}$,

$$\sum_{a=1}^{k} \tilde{\phi}_{\pi_a(\mathbf{v})}(\mathbf{v}) \geq \sum_{a=1}^{k} \phi_{\pi_a(\mathbf{v})}(\mathbf{v}),\ \forall k = 1, ..., n-1, \tag{2}$$

with the inequality being strict for some $k$, for a positive measure of $\mathbf{v}$'s.

Our welfare notion concerns the scope of efficiency, measured by the subset of schools that are efficiently allocated. To this end, fix an assignment $\mathbf{x} \in \Delta$ and a subset $K \subset S$ of schools. An assignment $\tilde{\mathbf{x}} \in \Delta$ is said to be a *within-$K$ reassignment of* $\mathbf{x}$ if $\tilde{x}_b = x_b$ for each $s_b \in S \backslash K$, and let $\Delta_{\mathbf{x}}^K \subset \Delta$ be the set of all such reassignments. Then, a *within-$K$ reallocation of* an allocation $\phi \in \mathcal{X}$ is an element of a set

$$\mathcal{X}_{\phi}^K := \{\tilde{\phi} \in \mathcal{X} | \tilde{\phi}(\mathbf{v}) \in \Delta_{\phi(\mathbf{v})}^K,\ \text{ a.e. } \mathbf{v} \in \mathcal{V}\}.$$

In words, a within-$K$ reallocation represents an outcome of students trading their shares of schools only within $K$.

**Definition 1.** *(i) For any $K \subset S$, an allocation $\phi \in \mathcal{X}$ is* ***Pareto efficient (PE) within*** *$K$ if there is no within-$K$ reallocation of $\phi$ that Pareto dominates $\phi$. (ii) For any $K \subset S$, an allocation $\phi \in \mathcal{X}$ is* ***ordinally efficient (OE) within*** *$K$ if there is no within-$K$ reallocation of $\phi$ that ordinally dominates $\phi$. (iii) An allocation is* ***PE*** *(resp.* ***OE****) if an allocation is PE (resp. OE) within $S$. (iv) An allocation is* ***pairwise PE*** *(resp.* ***pairwise OE****) if it is PE (resp. OE) within every $K \subset S$ with $|K| = 2$.*

These welfare criteria are quite intuitive. Suppose the students are initially endowed with ex ante shares $\phi$ of schools, and they can trade these shares among them. Can they trade mutually beneficially if the trading is restricted to the shares of $K$? The answer is no if allocation $\phi$ is *PE within $K$*. In other words, the size of the latter set represents the restriction on the trading technologies and thus determines the scope of markets within which efficiency is realized. The bigger this set is, the less restricted the agents are in realizing the gains from trade, so the more efficient the allocation is. Clearly, if an allocation is Pareto efficient within the set of all schools, then it is fully Pareto efficient. In this sense, we can view the size of such a set as a measure of efficiency.

A similar intuition holds with respect to ordinal efficiency. In particular, ordinal efficiency can be characterized by the inability to form a cycle of traders who can beneficially swap their probability shares of schools. Formally, let $\rhd^{\phi}$ be the binary relation on $S$ defined by

$$s_a \rhd^{\phi} s_b \Longleftrightarrow \exists A \subset \mathcal{V}, \mu(A) > 0, \text{ s.t. } v_a > v_b \text{ and } \phi_b(\mathbf{v}) > 0, \forall \mathbf{v} \in A;$$

that is, if a positive measure students prefer $s_a$ to $s_b$ but are assigned to $s_b$ with positive probabilities. We say that $\phi$ *admits a trading cycle within* $K$ if there exist $s_1, s_2, ...s_l \in K$ such that $s_1 \rhd^{\phi} s_2, ..., s_{l-1} \rhd^{\phi} s_l$, and $s_l \rhd^{\phi} s_1$. The next lemma is adapted from Bogomolnaia and Moulin (2001).

**Lemma 1.** *An allocation $\phi$ is OE within $K \subset S$ if and only if $\phi$ does not admit a trading cycle within $K$.*

Before proceeding further, we observe how different notions relate to one another.

**Lemma 2.** *(i) If an allocation is PE (resp. OE) within $K'$, then it is PE (resp. OE) within $K \subset K'$; (ii) An allocation is OE within $K \subset S$ if it is PE within $K$; (iii) If an allocation is OE within $K$ for any $K$ with $|K| = 2$, then it is PE within $K$; (iv) If an allocation is OE, then it is pairwise PE.*

Part (i) follows since a Pareto improving within-$K$ reallocation constitutes a Pareto improving within-$K'$ reallocation for any $K' \supset K$. Likewise, a trading cycle within any set forms a trading cycle within its superset. Part (ii) follows since if an allocation is not ordinally efficient within $K$, then it must admit a trading cycle within $K$, which produces a Pareto improving reallocation. Part (iii) follows since, whenever there exists an allocation that is not Pareto efficient within a pair of schools, one can construct a trading cycle involving two agents who would benefit from swapping their probability shares of these schools. Part (iv) then follows from Part (iii).

These characterizations are tight. The converse of Part (iii) does not hold for any $K$ with $|K| > 2$. In the example from the introduction, the DA allocation is OE but not PE. Likewise, an allocation that is PE within $K$ need not be OE within any $K' \supsetneq K$, since an allocation could be Pareto improved upon only via a trading cycle that includes a school in $K' \backslash K$. In that case, the allocation may be PE within $K$, yet it will not be OE within $K'$.

## 3.3 Alternative School Choice Procedures

We consider three alternative procedures for assigning students to the schools: (1) Deferred Acceptance with Single Tie-breaking (DA-STB), (2) Deferred Acceptance with Multiple Tie-Breaking (DA-MTB), and (3) Choice-Augmented Deferred Acceptance (CADA).

The alternative procedures differ only by the way the schools break ties. The tie-breaking rule is well-defined for DA-STB and DA-MTB, and it follows Step 2 of Section 2 in the case of CADA. These rules can be extended to the continuum of students in a natural way. The formal descriptions are provided in the Supplementary Notes; here we offer the following heuristic descriptions:

- **DA-STB:** Each student draws a number $\omega \in [0, 1]$ at random. A student with a lower number has a higher priority at *every* school than does a student with a higher number.

- **DA-MTB:** Each student draws $n$ independent random numbers $(\omega_1, ..., \omega_n)$ from $[0, 1]^n$. The $a$-th component, $\omega_a$, of student's random draw then determines her priority at school $s_a$, with a lower number having a higher priority than does a higher number.

- **CADA:** The mechanism draws two random numbers $(\omega_T, \omega_R) \in [0, 1]^2$ for each student. School $a$ then ranks those students who targeted that school, based on their values of $\omega_T$, and then ranks the others based on the values of $1 + \omega_R$ (with a lower number having a higher priority in both cases). In other words, those who didn't target the school receive a penalty score of 1.

For each procedure, the DA algorithm is readily defined using the appropriate tie-breaker and the students' ordinal preferences as inputs. The supplementary notes provide a precise algorithm, which is sketched here. At the first step, each student applies to her most preferred school. Every school $a$ tentatively admits up to unit mass from its applicants according to its priority order, and rejects the rest if there are any. In general, each student who was rejected in the previous step applies to her next preferred school. Each school considers the set of students it has tentatively admitted and the new applicants. It tentatively admits up to unit mass from these students in the order of its priority, and rejects the rest. The process converges when the set of students that are rejected has zero measure. Although this process might not complete in finite time, it converges in limit and the allocation in the limit is well defined. We focus on that limiting allocation.

More importantly, all three procedures are ordinally strategy proof:

**Theorem 1.** (ORDINAL STRATEGY-PROOFNESS) *In each of the three procedures, it is a (weakly) dominant strategy for each student to submit her ordinal preferences truthfully.*

*Proof:* The proof is in the supplementary notes.

## 3.4 Characterization of Equilibria

□ *DA-STB* and *DA-MTB*

In either form of DA algorithm, the resulting allocation is conveniently characterized by the "cutoff" of each school — namely the highest lottery number a student can have to get into that school. Specifically, the DA-STB process induces a cutoff $c_a \in [0,1]$ for each school $a$ such that a student who ever applies to school $a$ gets admitted by that school if and only if her (single) draw $\omega$ is less than $c_a$. We first establish that these cutoffs are well defined and generically distinct.

**Lemma 3.** *DA-STB admits a unique set of cutoffs $\{c_a\}_{a\in A}$ for the schools under DA-STB. Each cutoff is strictly positive and one of them equals 1. For a generic $\mathfrak{m}$, the cutoffs are all distinct.*

Importantly, these cutoffs pin down the allocation of all students. To see this, consider any student with $\mathbf{v}$ and a school $a$ with cutoff $c_a$. Suppose school $b$ has the highest cutoff among those schools that are preferred to $s_a$ by that student. If the cutoff of school $b$ has $c_b > c_a$, then the student will never get assigned to school $a$ since whenever she has a draw $\omega < c_a$ (good enough for $s_a$), she will get into school $b$ or better. If $c_b < c_a$, however, then she will get into school $a$ if and only if she receives a draw $\omega \in [c_b, c_a]$. The probability of this event is precisely the distance between the two cutoffs, $c_a - c_b$. Formally, let $S(a,\mathbf{v}) := \{s_b \in S | v_b > v_a\}$ denote the set of schools more preferred to $i$ by type-$\mathbf{v}$ students. Then, the allocation $\phi^S$ arising from DA-STB is given by

$$\phi_a^S(\mathbf{v}) := \max\{c_a - \max_{s_b\in S(a,\mathbf{v})} c_b, 0\}, \forall \mathbf{v}, \forall a \in A,$$

where $c_\emptyset := 0$.

DA-MTB is similar to DA-STB, except that each student has independent draws $(\omega_1, ..., \omega_n)$, one for each school. The DA process again induces a cutoff $\tilde{c}_a \in [0,1]$ for each school $a$ such that a student who ever applies to school $a$ gets assigned to it if and only if her draw for school $a$, $\omega_a$, is less than $\tilde{c}_a$. These cutoffs are well defined.

**Lemma 4.** *DA-MTB admits a unique set of cutoffs $\{\tilde{c}_a\}_{a\in A}$. Each cutoff is strictly positive and one of them equals 1. For a generic $\mathfrak{m}$, the cutoffs are all distinct.*

Given the cutoffs $\{\tilde{c}_a\}_{a\in A}$, a type $\mathbf{v}$-student receives school $a$ whenever she has a rejectible draw $\omega_b > \tilde{c}_b$ for each $s_b \in S(a,\mathbf{v})$ she prefers to school $a$ and when she has an acceptable draw $\omega_a < \tilde{c}_a$ for school $a$. Formally, the allocation $\phi^M$ from DA-MTB is determined by:

$$\phi_a^M(\mathbf{v}) := \tilde{c}_a \prod_{s_b\in S(a,\mathbf{v})} (1 - \tilde{c}_b), \forall \mathbf{v}, \forall a \in A,$$

with the convention $\tilde{c}_\emptyset := 0$.

□ *CADA*

As with the two other procedures, given the students' strategies on their messages, the DA process induces cutoffs for the schools, one for each school in $[0, 2]$. Of particular interest is the equilibrium in the students' choices of messages. Given Theorem 1, the only nontrivial part of the students' strategy concerns her "auxiliary message." Let $\sigma = (\sigma_1, ..., \sigma_n) : \mathcal{V} \mapsto \Delta$ denote the students' mixed strategy, whereby a student with $\mathbf{v}$ targets $s_a$ with probability $\sigma_a(\mathbf{v})$. We first establish existence of equilibrium.

**Theorem 2.** (EXISTENCE) *There exists an equilibrium $\sigma^*$ in pure strategies.*

We say that a student *applies to school $a$* if she is rejected by all schools she lists ahead of $a$ in her (truthful) ordinal list. We say that a student *subscribes to school $a \in S$* if she targets school $a$ *and* applies to that school during the DA process. (The latter event depends on where she lists school $a$ in her ordinal list and the other students' strategies, as well as the outcome of tie breaking). Let $\bar{\sigma}^*_a(\mathbf{v})$ be the probability that a student $\mathbf{v}$ subscribes to school $a$ in equilibrium. We say a school $a \in S$ is *oversubscribed* if $\int \bar{\sigma}^*_a(\mathbf{v}) d\mu(\mathbf{v}) \geq 1$ and *undersubscribed* if $\int \bar{\sigma}^*_a(\mathbf{v}) d\mu(\mathbf{v}) < 1$. In equilibrium, there will be at least (generically, exactly) one undersubscribed school which anybody can get admitted to (that is, even when she fails to get into any other schools she listed ahead of that school). Formally, a school $w \in S$ is said to be "worst" if its cutoff on $[0, 2]$ equals precisely 2. Then, we have the following lemma.

**Lemma 5.** *(i) Any student who prefers the worst school the most is assigned to that school with probability 1 in equilibrium. (ii) If her most preferred school is undersubscribed but not the worst school, then she targets that school in equilibrium. (iii) For almost every student with $\mathbf{v}$ such that $\pi_1(\mathbf{v}) \neq w$, $\sigma^*(\mathbf{v}) = \bar{\sigma}^*(\mathbf{v})$ in equilibrium.*

In light of Lemma 5-(iii), we shall refer to "targets a school $a$" simply as "subscribes to school $a$."

# 4 Welfare Analysis of Alternative Procedures

It is useful to begin with an example. Suppose there are three schools, $S = \{a, b, c\}$, and three types of students $\mathcal{V} = \{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3\}$, each with $\mu(\mathbf{v}^i) = 1$, and their vNM values are described as follows.

| | $v^1_j$ | $v^2_j$ | $v^3_j$ |
|---|---|---|---|
| $j = a$ | 5 | 4 | 1 |
| $j = b$ | 1 | 2 | 5 |
| $j = c$ | 0 | 0 | 0 |

Consider first DA-MTB. Each student draws three lottery numbers, $(\omega_a, \omega_b, \omega_c)$, one for each school. The schools $a$, $b$ and $c$ then have cutoffs $\tilde{c}_a \approx 0.39, \tilde{c}_b \approx 0.45$, and $\tilde{c}_c = 1$, respectively. The resulting allocation is $\phi^M(\mathbf{v}^1) = \phi^M(\mathbf{v}^2) \approx (0.39, 0.27, 0.33)$ and $\phi^M(\mathbf{v}^3) \approx (0.22, 0.45, 0.33)$. This allocation is PE within $\{a, c\}$ and within $\{b, c\}$, but not even OE within $\{a, b\}$. The ordinal inefficiency within $\{a, b\}$ can be seen by the fact that type-$\{\mathbf{v}^1, \mathbf{v}^2\}$ students have positive shares of school $b$, and type-$\mathbf{v}^3$ students have positive share of school $a$, which they can swap with each other to do better. This feature originates from the independent drawings of priority lists for the schools. For instance, as in the figure, type-$\{\mathbf{v}^1, \mathbf{v}^2\}$ students may draw $(\omega_a, \omega_b)$ and type-$\mathbf{v}^3$ students may draw $(\omega'_a, \omega'_b)$. Hence, we have $a \rhd^{\phi^M} b \rhd^{\phi^M} a$. (Note that the cutoff for school $c$ is 1, which explains why the allocation is PE within $\{a, c\}$ and within $\{b, c\}$.)

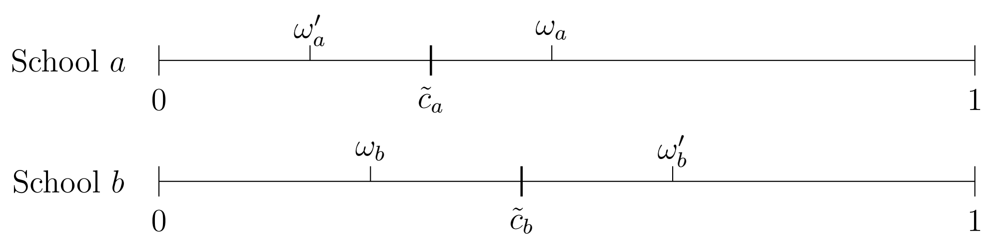


Figure 1: Ordinal inefficiency within $\{a, b\}$ under DA-MTB.

DA-STB avoids this problem, since each student draws only one lottery number for all schools. In this example, the cutoffs of schools $a, b$ and $c$ are $c_a = 1/2$, $c_b = 2/3$, and $c_c = 1$, respectively. The resulting allocation is $\phi^S(\mathbf{v}^1) = \phi^S(\mathbf{v}^2) = (\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$ and $\phi^S(\mathbf{v}^3) = (0, \frac{2}{3}, \frac{1}{3})$. This allocation is OE, and thus pairwise PE (by Lemma 2).


Figure 2: Ordinal efficiency of DA-STB

To see this, consider any students who strictly prefer school $b$ to school $a$. In our example, type-$\mathbf{v}^3$ students have such preference. These students can never be assigned to school $a$ since, whenever they have draws acceptable for school $a$ (for instance $\omega < c_a$ in Figure 2), they will choose school $b$ and admitted by it. Hence, we cannot have $b \rhd^{\phi^S} a$. A similar logic implies that we cannot have $c \rhd^{\phi^S} b$.[9] Hence, $\phi^S$ admits no trading cycle. Despite the superiority over DA-MTB, the DA-STB allocation is not fully PE; type-$\mathbf{v}^1$ students can profitably trade with

[9] In this example, no student prefers school $c$ to $b$, but the logic applies even if there were students with such a preference.

type-$\mathbf{v}^2$ students, selling probability shares of schools $a$ and $c$ in exchange for probability share of school $b$.

Consider lastly CADA. As with the two DA mechanisms, all students rank the schools truthfully; and type-$\{\mathbf{v}^1, \mathbf{v}^2\}$ students target school $a$ and type-$\mathbf{v}^3$ target school $b$. The resulting equilibrium allocation is $\phi^*(\mathbf{v}^1) = \phi^*(\mathbf{v}^2) = (\frac{1}{2}, 0, \frac{1}{2})$ and $\phi^*(\mathbf{v}^3) = (0, 1, 0)$. Notice that no type-$\{\mathbf{v}^1, \mathbf{v}^2\}$ students are ever assigned to school $b$, which means in this case the allocation is fully PE.

These observations are generalized as follows:

**Theorem 3.** (DA-MTB) *(i) The allocation $\phi^M$ from DA-MTB is PE within $\{a, w\}$ for each $s_a \in S \setminus \{w\}$. (ii) Generically, there exists no $K \subset S$ with either $|K| > 2$ or $|K| = 2$ but $\tilde{c}_b < 1, \forall s_b \in K$ such that $\phi^M$ is OE within $K$.*

**Theorem 4.** (DA-STB) *(i) The allocation $\phi^S$ from DA-STB is OE and is thus pairwise PE. (ii) For a generic $\mathfrak{m}$, there exists no $K \subset S$ with $|K| > 2$ such that $\phi^S$ is PE within $K$.*

In sum, DA-STB can yield an ordinally efficient allocation in the large economy, but this is the most that can be expected from DA-STB, in the sense that the scope of efficiency is generically limited to (sets of) three schools.

**Remark 1.** *With finite students, the allocation from DA-STB is ex post Pareto efficient but is not OE. But as the number of students and school seats get large, the DA-STB allocation becomes OE in the limit. This is an implication of Che and Kojima (2010), who show that the random priority rule (which coincides with our DA-STB) becomes indistinguishable from the probabilistic serial mechanism (which is known to be OE) as the economy grows large. When the schools have intrinsic priorities, the DA-STB is not even ex post Pareto efficient (Abdulkadiroglu, Pathak and Roth (forthcoming)).*

**Theorem 5.** (CADA) *(i) An equilibrium allocation $\phi^*$ of CADA is OE and is thus pairwise PE. (ii) An equilibrium allocation of CADA is PE within the set of oversubscribed schools. (iii) If all but one school is oversubscribed, then the equilibrium allocation of CADA is PE.*

Theorem 5-(ii) and (iii) showcase the ex ante efficiency benefit associated with CADA. The benefit parallels that of a competitive market. Essentially, CADA activates "competitive markets" for oversubscribed schools. Each student is given a "budget" of unit probability she can allocate across alternative schools for targeting. A given unit probability can buy different numbers of shares for different schools, depending on how many others name those schools. If a mass $z_a \geq 1$ students applies to school $a$, allocating a unit budget can only buy a share $1/z_a$. Hence, the relative congestion at alternative schools, or their relative popularity, serves as relative "prices" for these schools.[10] In a large economy, individual students take these prices as given, so the prices play the usual role of allocating resources efficiently. It is therefore not surprising that the proof follows the First Welfare Theorem.

Why are competitive markets limited only to oversubscribed schools? Why not undersubscribed schools? Recall that one can get into an undersubscribed school in two different ways: she can target it, in which case she gets assigned to it for sure if she applies to it. Alternatively, she can target an oversubscribed school but the school may reject her, in which case she may still get assigned to that undersubscribed school via the usual DA channel. Clearly, assignment via this latter channel does not respond to, or reflect, the "prices" set by the targeting behavior. Consequently, competitive markets do not extend to the undersubscribed schools.

Finally, Part (i) asserts ordinal efficiency for CADA. At first glance, this feature may be a little surprising in light of the fact that different schools use different priority lists. As is clear from DA-MTB, this feature is susceptible to ordinal inefficiency. The CADA equilibrium is OE, however. To see this, observe first that any student who is assigned to an oversubscribed school with positive probability must strictly prefer it to any undersubscribed school (or else she should have secured assignment to the latter school by targeting it). Thus, we cannot have $b \rhd^{\phi^*} a$ if school $b$ is undersubscribed and school $a$ is oversubscribed. This means that if the allocation admits any trading cycle, it must be within oversubscribed schools or within undersubscribed schools. The former is ruled out by Part (ii) and the latter by the same argument as Theorem 4-(i).

The characterization of Theorem 5-(ii) is tight in the sense that there is generally no bigger set that includes all oversubscribed schools *and some undersubscribed school* that supports Pareto efficiency (see the supplementary notes).

Theorem 5 refers to an endogenous property of an equilibrium, namely the set of over/undersubscribed schools. We provide a sufficient condition for this property. For each school $a \in S$, let $m^*_a := \mu(\{\mathbf{v} \in \mathcal{V} | \pi_1(\mathbf{v}) = a\})$ be the measure of students who prefer $s_a$ the most. We then say a school $a$ is **popular** if $m^*_a \geq 1$, namely, the size of the students whose most preferred school is $s_a$ is as large as its capacity.

It is easy to see that every popular school must be oversubscribed in equilibrium. Suppose to the contrary that a popular school $a$ is undersubscribed. Then, by Lemma 5-(ii), every student

[10] Given the DA format, a student may be assigned to an undersubscribed school after targeting (and failing to get into) an oversubscribed school. This may cause a potential spill-over from consumption of an oversubscribed school toward undersubscribed schools. This spill-over does not undermine the efficient allocation, however. Under our CADA procedure, targeting alternative oversubscribed schools have no impact on the conditional probability of assignment with undersubscribed schools, since the tie breaking at non-target schools are determined by a *separate* random priority lists (see footnote 8).

with $\mathbf{v}$ with $\pi_1(\mathbf{v}) = a$ must subscribe to $s_a$, a contradiction. Since every popular school is oversubscribed, the next result follows from Theorem 5.

**Corollary 1.** *Any equilibrium allocation of CADA is PE within the set of popular schools.*

It is worth emphasizing that the popularity of a school is sufficient, but not necessary, for that school to be oversubscribed. In many situations, many non-popular schools will be oversubscribed. In this sense, Corollary 1 understates the benefit of CADA. This point is confirmed by our simulation in Section 5. Even though full PE is not ensured by any of the three mechanisms, our results imply the following comparison between standard DA and CADA.

**Corollary 2.** *(i) If $n \geq 3$, generically the allocations from DA-STB and DA-MTB are not PE. (ii) The equilibrium allocation of CADA is PE if all but one school is popular.*

The results so far give a sense of a three-way ranking of DA-MTB, DA-STB, and CADA. Specifically, if the allocation from DA-MTB is PE within $K \subset S$, then so is the allocation from DA-STB, although the converse does not hold; and if the allocation from DA-STB is PE within $K' \subset S$, then so is the allocation from CADA, although the converse does not hold. Between the two DA algorithms, the DA-STB allocation is OE, whereas the DA-MTB allocation is not pairwise PE.

In particular, the CADA allocation is PE within a strictly bigger set of schools than the allocations from DA algorithms, if there are more than two popular schools. Unfortunately, this is not the case when all students have the same ordinal preference. This case, though special, is important since parents often tend to rank schools similarly. In this case, there is only one popular school in a CADA equilibrium, so Theorem 5 and Corollary 1-(i) do little to distinguish CADA from DA-STB. Nevertheless, we can find the CADA to be superior in a more direct way. To this end, let $\mathcal{V}^U := \{\mathbf{v} \in \mathcal{V} | v_1 > ... > v_n\}$.

**Theorem 6.** *Suppose all students have the same ordinal preferences in the sense $\mu(\mathcal{V}^U) = \mu(\mathcal{V})$. The equilibrium allocation of CADA (weakly) Pareto dominates the allocation arising from DA-STB and DA-MTB.*

This result generalizes the example discussed in the introduction. If all students have the same ordinal preferences, the DA algorithm with any random tie-breaking treat all students in the same way, meaning that each student is assigned to each school with equal probability. Under CADA, the students can at least replicate this random assignment via targeting.

# 5 Simulations

The theoretical results in the previous sections do not speak to the magnitude of efficiency gains or losses achieved by each mechanism. Here, we provide a numerical analysis of the magnitude via simulations.

In our numerical model, we have 5 schools, each with 20 seats and 100 students. Student $i$'s vNM value for school $j$, $\tilde{v}_{ij}$, is given by

$$\tilde{v}_{ij} = \alpha u_j + (1-\alpha) u_{ij}$$

where $\alpha \in [0,1]$, $u_j$ is common across students and $u_{ij}$ is specific to student $i$ and school $j$. For each $\alpha$, we draw $\{u_j\}$ and $\{u_{ij}\}$ uniformly and independently from the interval $[0,1]$ to construct student preferences. We then normalize each student's vNM utilities by $v_{ij} = \zeta_j(\tilde{\mathbf{v}}_i) := \frac{\tilde{v}_{ij} - \min_{j'} \tilde{v}_{ij'}}{\max_{j'} \tilde{v}_{ij'} - \min_{j'} \tilde{v}_{ij'}}$. Under this normalization, the values of schools range from zero to one, with the value of the least preferred school set to zero and that of the most preferred to one. This normalization is invariant to affine transformation in the sense that $\zeta_j(\tilde{v}_{is_1}, ..., \tilde{v}_{is_5}) = \zeta_j(\alpha\tilde{v}_{is_1} + \beta, ..., \alpha\tilde{v}_{is_5} + \beta)$, for any $\alpha \in \mathbb{R}_{++}, \beta \in \mathbb{R}$.

The students' preferences become similar to one another both ordinally and cardinally as $\alpha$ gets large. In one extreme case with $\alpha = 0$, students' preferences are completely uncorrelated; in the other with $\alpha = 1$, students have the same cardinal (as well as ordinal) preferences. Given a profile of normalized vNM utility values, we simulate DA-STB and DA-MTB, compute a complete-information Nash equilibrium of CADA and the resulting CADA allocation. We repeat this computation 100 times each with a new set of (randomly drawn) vNM utility values for all values of $\alpha$. In addition, we solve for a first-best solution, which is the utilitarian maximum for each set of vNM utility values. We then compute the average welfare under each mechanism, i.e., the total expected utilities realized under a given mechanism averaged over 100 draws (see the supplementary notes for details).

In Figure 3, we compare the three mechanisms to the first best solution. We plot the welfare of each mechanism as the percentage of the welfare of the first best solution. Two observations emerge from this figure. First, the welfare generated by each mechanism follows a U-shaped pattern. Second, CADA outperforms DA-STB, which in turn outperforms DA-MTB at every value of $\alpha$, and the gap in performance between CADA and the other mechanisms grows with $\alpha$. All three mechanisms perform almost equally well and produce about 96% of the first-best welfare when $\alpha = 0$. In this case, students have virtually no conflicts of interests, and each mechanism more or less assigns students to their first choice schools. The welfare gain of CADA increases as $\alpha$ increases. This is due to the fact that competition for one's first choice increases as $\alpha$ increases (and students' ordinal preferences get similar to one another). In those instances, who gets her first choice matters. While DA-STB and DA-MTB determine this purely randomly, CADA does so based on students' messages. Intuitively, if a student's vNM value for a school increases, the likelihood of the student targeting that school in an equilibrium of CADA — therefore the likelihood of her getting into that school — increases. This feature of CADA contributes to its welfare gain. DA-STB and DA-MTB start catching up with CADA at $\alpha = 0.9$. In this case, students have almost the same cardinal preferences, so any matching is close to being ex ante efficient. At $\alpha = 0.9$, CADA achieves 95.5% of the first best welfare, whereas DA-STB achieves 92.2%.[11]

Figure 4 gives further insight into the workings of the mechanisms. It shows the percentage of students getting their first choices under each mechanism. First, DA-MTB assigns noticeably smaller numbers to first choices. This is due to the artificial stability constraints created by multiple tie breaking, which also explains the bigger welfare loss associated with DA-MTB. The patterns for CADA and DA-STB are more revealing. In particular, both assign almost the same number of students to their first choices for each value of $\alpha$. That is, whereas the poor welfare performance of DA-MTB is explained by the low number of students getting their first choices, the difference between the other two is explainable not by *how many* students, but rather by *which students* are assigned to their first choices.

This is illustrated more clearly by Figure 5, which shows the ratio of the mean utility of those who get their $k$-th choice under CADA to the mean utility of those who get their $k$-th choice under DA-STB at the realized matchings, for $k = 1, 2, 3$. Specifically, those who get their $k$-th choice achieve a higher utility under CADA than under DA-STB for each $k = 1, 2, 3$. The utility gain is particularly more pronounced for the receivers of their second or third choices. This simply reflects the feature of CADA that assigns students based on their preference intensities: under CADA, those who have less to lose from the second- or third-best choices are more likely to target those schools, and are thus more likely to compose such assignments.

Figure 6 shows that the number of oversubscribed schools is larger on average than the number of popular schools. Note that the average number of oversubscribed schools is larger than 2 at all values of $\alpha$. Recalling our Theorems 4 and 5, DA-STB is never PE within a set of more than 2 schools, whereas CADA is PE within the set of oversubscribed schools. Figure 6 thus shows the scope of efficiency achieved by CADA can be much higher than is predicted by Corollary 1. It is also worth noting that the average number of oversubscribed schools exceeds 3 for $\alpha \leq 0.4$. This implies that there are often 4 oversubscribed schools. At those instances, CADA achieves full Pareto efficiency (recall Theorem 5-(ii)).

[11] At the extreme case of $\alpha = 1$, preferences are the same so every matching is efficient and the welfare generated by each mechanism is equal to the first best welfare.

In practice, some schools have (non-strict) intrinsic priorities. We thus study their impact on assignments numerically. To this end, we modify our model as follows: Each school has two priority classes, high priority and low priority. For each preference profile above, we assume that 50 students have high priority in their first choice and low priority in their other choices, 30 students have high priority in their second choice and low priority in their other choices, and 20 students have high priority in their third choice and low priority in their other choices.[12]

It is well known that standard mechanisms such as DA do not produce student optimal stable matching in the presence of school priorities. Erdil and Ergin (2008) have proposed a way to attain constrained ex post efficiency subject to respecting school priorities, via performing so-called stable improvement cycles after an initial DA assignment. We thus simulate this algorithm, referred to as DASTB+SIC, to see how it compares with the CADA.

In Figure 7, we compare CADA, DA-STB and DA-STB+SIC again measured as percentage of first-best welfare. Again, CADA outperforms DA-STB for all values of $\alpha$. Since DA-STB+SIC is designed to achieve constrained ex post efficiency (while CADA and DA-STB are not), it is not surprising that the former does better when $\alpha$ is relatively small. In that case, students' ordinal preferences are sufficiently dissimilar that ordinal efficiency matters. As $\alpha$ gets large, however, ordinal efficiency becomes less relevant and cardinal efficiency becomes more important. For $\alpha \geq 0.5$, CADA catches up with DA-STB+SIC and outperforms it as $\alpha$ gets large. In particular, when $\alpha$ is close to 1, virtually all matchings are ex post efficient, so DA-STB+SIC has little bite. The cardinal efficiency still matters, and in this regard, CADA does better than the other mechanisms. This finding is noteworthy since parents are likely to have similar ordinal preferences in real-life choice settings. In those instances, CADA allocates schools more efficiently than other mechanisms in ex ante welfare.

# 6 Discussion

## 6.1 Enriching the Auxiliary Message

CADA can be modified to allow for more complicated auxiliary messages, perhaps at the expense of some practicality. For instance, the auxiliary message can include a rank order of schools up to $k \leq n$, with a tie broken in the lexicographic fashion according to this rank order: students targeting a school at a higher lexicographic component is favored by that school in a tie relative to those who do not target or target it at a lower lexicographic component. We call the associated CADA a ***CADA of degree*** $k$.

[12]This assumption is in line with the stylized fact about the Boston school system.

It is worth noting that the CADA of degree $n$ coincides with the Boston mechanism if the schools have no priorities and if all students have the same ordinal preferences. Such an enriching of the auxiliary message does not alter the qualitative features of CADA. In particular, an argument analogous to that of Theorem 6 applies to CADA of any degree, which has a rather surprising implication:

**Theorem 7.** *If all students have the same ordinal preferences and the schools have no priorities, then the Boston mechanism weakly Pareto dominates the DA algorithm.*

A richer message space could allow students to signal their relative preference intensities better, and this may lead to a better outcome (see Abdulkadiroglu, Che and Yasuda (2008) for an example). A richer message space need not deliver a better outcome, however. With more messages, students have more opportunities to express their relative preference intensities over different sets of schools. The increased opportunities may act as substitutes and militate each other. For instance, an increased incentive to self select at low-tier schools may lessen a student's incentive to self select at high-tier schools. This kind of "crowding out" arises in the next example.

**Example 1.** *There are 4 schools, $S = \{a, b, c, d\}$, and two types of students $\mathcal{V} = \{\mathbf{v}^1, \mathbf{v}^2\}$, with $\mu(\mathbf{v}^1) = 3$ and $\mu(\mathbf{v}^2) = 1$.*

| | $v_j^1$ | $v_j^2$ |
|---|---|---|
| $j = a$ | 12 | 8 |
| $j = b$ | 2 | 4 |
| $j = c$ | 1 | 3 |
| $j = d$ | 0 | 0 |

*Consider first CADA of degree 1. Here, type-$\mathbf{v}^1$ students target $a$, and type-$\mathbf{v}^2$ students target school $b$. In other words, the latter type of students self select into the second popular school. The resulting allocation is $\phi^*(\mathbf{v}^1) = (\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3})$ and $\phi^*(\mathbf{v}^2) = (0, 1, 0, 0)$. The expected utilities are $EU^1 = 4.33$ and $EU^2 = 4$. In fact, this allocation is PE.*

*Suppose now CADA of degree 2 is used. In equilibrium, type-$\mathbf{v}^1$ students choose school $a$ and $b$ as their first and second targets, respectively. Meanwhile, type-$\mathbf{v}^2$ students choose school $a$ (instead of school $b$!) for their first target and school $b$ for their second target. Here, the opportunity for type 2 students to self select at a lower-tier school (school $c$) blunts their incentive to self select at a higher-tier school (school $b$). The resulting allocation is thus $\phi^{**}(\mathbf{v}^1) = (\frac{1}{4}, \frac{1}{3}, \frac{1}{12}, \frac{1}{3})$ and $\phi^{**}(\mathbf{v}^2) = (\frac{1}{4}, 0, \frac{3}{4}, 0)$, which yield expected utilities of $\overline{EU}^1 = 3.75$ and $\overline{EU}^2 = 4.25$. This allocation is not PE since type-$\mathbf{v}^2$ students can trade probability shares of school $a$ and $c$ in exchange for probability share of $b$, with type-$\mathbf{v}^1$ students.*

*Even though $\phi^*$ does not Pareto dominate $\phi^{**}$, the former is PE whereas the latter is not. Further, the former is superior to the latter in the Utilitarian sense (recall that students' payoffs are normalized so that they aggregate to the same value for both types): the former gives aggregate utilities of* 17*, the highest possible level, whereas the latter gives* 15.5*.*

This example suggests that the benefit from enriching the message space is not unambiguous. This is a potentially important point. In practice, expanding a message space adds a burden on the parents to be more strategically sophisticated. Hence avoiding such a demand for strategic sophistication is an important quality for a procedure to succeed. This adds to the appeal of the simple CADA (i.e., of degree 1).

## 6.2 Strategic Naivety

Since CADA involves some "gaming" aspect, albeit limited to tie-breaking, a natural concern is that not all families may be strategically competent. This concern has arisen in the context of the Boston mechanism. It has been noted that some significant percentage of families have played suboptimal strategies, for instance, wasting their second top choices to schools that they could only get into by top-ranking them (Abdulkadiroglu, Pathak, Roth and Sönmez, 2006). Such mistakes may arise because of the lack of knowledge about how the system works or about which schools are popular. The same concern may arise with respect to CADA, in that some families may not understand well the role the auxiliary message plays in the system and/or they may not judge accurately how over/undersubscribed various schools will turn out to be.

It is thus important to investigate how CADA will perform when some families are not strategically sophisticated. To this end, we consider students who are "naive" in the sense that they always target their most preferred schools in the auxiliary message. Targeting the most preferred school appears to be a simple, but reasonable, choice when a student is unsure about the popularity of alternative schools or is unclear about the role the auxiliary message plays in the assignment. Such a strategy will indeed be a best response for many situations, particularly if the first choice is distinctively better than the rest of the choices, so it could be a reasonable approximation of "naive" behavior. We assume that there is a positive measure of students who are naive in this way, and the others know the presence of these students and their behavior, and respond optimally against them. Surprisingly, the presence of naive students do not affect the main welfare results in a qualitative way.

**Theorem 8.** *In the presence of naive students, the equilibrium allocation of CADA satisfies the following properties: (i) The allocation is OE, and is thus pairwise PE. (ii) The allocation is*

*PE within the set $K$ of oversubscribed schools. (iii) If every student is naive, then the allocation is PE within $K \cup \{l\}$ for any undersubscribed school $l \in J := S \backslash K$.*

Theorem 8-(i) and (ii) are qualitatively the same as the corresponding parts of Theorem 5. Further, Lemma 5-(ii) remains valid in the current context, implying that any popular schools must be oversubscribed here as well. Hence, the same conclusion as Corollary 1 holds.

**Corollary 3.** *In the presence of naive students, the equilibrium allocation of CADA is PE within the set of popular schools.*

## 6.3 CADA with "Safety Valve"

The preceding subsection has seen that the main welfare property of CADA extends to the situation where some students behave naively. This does not mean, however, that naive students are not disadvantaged by the others who behave more strategically.[13] The CADA mechanism can be modified to provide an extra safeguard for those who are averse to strategic aspect of the game. This can be done by augmenting the message space to include an "exit option." First, we can run DA for *all* students with a random tie breaking; those who invoke the exit option are then assigned based on the outcome. Next, those who have not invoked the exit option can be assigned based on their targeting. This modification yields an allocation that Pareto dominates the standard DA algorithms (see Abdulkadiroğlu, Che and Yasuda (2008) for detail).

## 6.4 Dynamic Implementation

As noted, the welfare benefit of CADA originates from the competitive markets it induces. Unlike the usual markets where there are explicit prices, however, in the CADA-generated markets, students' beliefs about the relative popularity of schools act as the prices. Hence, for the CADA to have the desirable welfare benefit, their beliefs must be reasonably accurate. In practice, students/parents' beliefs about schools are formed based on their reputations; thus, as long as the school reputations are stable, they can serve as reasonably good proxies for the prices. Nevertheless, the students may not share the same beliefs and the beliefs may not be accurate, in which case CADA procedure will not implement the CADA equilibrium precisely.

The CADA mechanism can be modified to implement the desired equilibrium more precisely. The idea is to allow students to dynamically revise their target choices based on the population

[13] Pathak and Sonmez (2008) analyze the Boston mechanism when some students are strategically naive, and show that, at a Pareto dominant equilibrium of the game, the sophisticated students benefit in comparison to the DA mechanism at the expense of the naive players.

distribution of choices (which is made public). By making their choices final only when the number of students changing their choices fall under a certain threshold, we can induce a best response dynamics, which will implement the desired equilibrium precisely whenever it converges (see Abdulkadiroğlu, Che and Yasuda (2008) for detail).

### 6.5 Excess Capacities and Outside Options

Thus far, we have made simplifying assumptions that the aggregate measure of students equal the aggregate capacities of public schools and that all students find each public school acceptable. These assumptions may not hold in reality. While public schools must guarantee seats to students, all the seats need not be filled. And some students may find outside options, such as home or private schooling, better than some public schools. One can relax these assumptions by letting the aggregate capacities to be (weakly) greater than $n$ and by endowing each student an outside option with value drawn from $[0, 1]$.[14] Extending the model in this way entails virtually no changes in the main tenet of our paper. All theoretical results continue to hold in this relaxed environment. A subtle difference arises since, with excess capacities, there may be more than one school with cutoff equal to one under DA-MTB, so its allocation may become PE within more pairs of schools. Nevertheless, Theorems 1-8 remain valid. For instance, the DA-STB allocation is ordinally efficient. The CADA allocation is ordinally efficient and Pareto efficient within oversubscribed, and thus popular, schools.

## 7 Conclusion

In this paper, we propose a new deferred acceptance procedure in which students are allowed, via signaling of their preferences, to influence how they are treated in a tie for a school. This new procedure, choice-augmented DA algorithm (CADA), makes the most of two existing procedures, the Gale-Shapely deferred acceptance algorithm (DA) and the Boston mechanism. While the DA achieves the strategyproofness, an important property in the design of school choice programs, it limits students' abilities to communicate their preference intensities, which entails an ex ante inefficient allocation when schools are indifferent among students with the same ordinal preferences. The Boston mechanism, on the other hand, is responsive to the agents' cardinal preferences and may achieve more efficient allocation than the DA, but fails to satisfy strategyproofness. We show that, by allowing students to influence tie-breaking via additional

[14]This modeling approach implicitly assumes the outside options to have unlimited capacities, which may not accurately reflect the scarcity of outside option such as private schooling.

communication, CADA implements a more efficient ex ante allocation than the standard DA algorithms, without sacrificing the strategyproofness of ordinal preferences.

There are alternative mechanisms that could also perform well in the cardinal welfare (or more precisely in ex ante welfare). First, the Boston mechanism allows students to express their cardinal preferences through the ordinal ranking of the schools. Under the Boston mechanism, each school assigns its seats according to the order students rank that school during registration; each school accepts first those who rank it first, accepts those who rank it second only when seats are available, and so forth. Under this procedure, a student's ranking of a school influences her odds of getting assigned a seat at that school, since a student who ranks the school low will only have a chance to get a seat after all those who ranked the school more highly are accommodated. This could entail better ex ante welfare than the DA.[15] But, this comes at the expense of strategy-proofness, however, for students may have the incentive to move up the ranking of less preferred but more viable option over the more preferred option with slim odds. Further, the participants may not be well coordinated in their strategic play, which may result in extra inefficiencies.[16] By contrast, CADA is strategy-proof in the ordinal preference rankings. While CADA involves strategic plays, its scope is limited to targeting, and its influence is kept within a priority class. In fact, targeting involves a relatively simple and straightforward strategic decision. We thus believe that the scope for miscoordination is limited in CADA. Further, we provide a dynamic implementation of targeting game that facilitates strategic coordination of the students.

Another mechanism that incorporates cardinal welfare is the pseudo-market mechanism proposed by Hylland and Zeckhauser (1979).[17] This mechanism purports to install competitive

[15]For the example discussed in Introduction, Boston mechanism also implements the desirable assignment. Students 1 and 2 have a dominant strategy of ranking the schools truthfully, and student 3 has a best response of (strategically) ranking school $b$ as her first choice. Consequently, student 3 is assigned $b$, and the other two are assigned between $a$ and $c$ with equal probability. (see Abdulkadiroğlu, Che and Yasuda, 2011). Indeed, Miralles (2008) applies the arguments developed in this paper to show that a variant of the Boston mechanism has a similar ex ante welfare property as the CADA. Elaborate.

[16]See Ergin and Sönmez (2006).

[17]A similar mechanism is also used in course allocation mechanisms (see Budish and Cantillon, 2011). Sönmez and Ünver (forth) imbed the DA algorithm in "course bidding" employed by some business schools. These two proposals differ in the application, however, as well as in the nature of the inquiry: we are interested in studying the benefit of adding a "signaling" element to the DA algorithm. By contrast, their interest is in studying the effect of adding ordinal preferences and the DA feature to course bidding.

In a broader sense, our paper is an exercise of mechanism design without monetary transfers, and in fact it is closer in nature to the recent ideas of "storable votes" (Casella, 2005) and "linking decisions" (Jackson and Sonnenschein, 2007). Just like them, CADA "links" how a student is treated in a tie at one school to how she is treated in a tie at another school, and this linking makes communication credible. Clearly, applying the idea in a markets for trading probability shares of alternative objects using a fictitious currency. Specifically, the mechanism endows each agent with a fixed budget in a fictitious currency, 100 tokens say, and allows the agents to spend their budget endowments to "buy" probability shares of alternative goods, and the price per unit probability of owning each good is then adjusted to clear the markets. For large markets, this mechanism admits a competitive equilibrium, which is ex ante efficient by the first welfare theorem.[18] Our result has a similar flavor. Indeed, the main contribution of our paper is to recognize that adding a signaling device as simple as targeting a school can have the same kind of "market-activating" effects as the pseudo-market mechanism. Although CADA does not generally attain full ex ante efficiency, the strategic environment is simple and the strategic deliberation required for the agents is not so demanding; by contrast, formulating ones' cardinal utilities (instead of simply "acting on them") could be more onerous, and the consequence of miscalculation on efficiency may be large. Most important, school priorities are already imbedded in CADA, whereas the pseudo-market mechanism does not include priorities. This is an important distinction, to the extent that priorities are an important feature of school choice.

# References


[1] Abdulkadiroğlu, A., Che, Y.-K., and Yasuda, Y. (2008), "Expanding "Choice" in School Choice," PER Discussion Papers 0809-09, Columbia University.

[2] Abdulkadiroğlu, A., Che, Y.-K., and Yasuda, Y. (2011), "Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered," *American Economic Review*, 101, 399-410.

[3] Abdulkadiroğlu, A., Pathak, P., and Roth, E. (2005), "The New York City High School Match," *American Economic Review*, Papers and Proceedings, 95, 364-367.

[4] Abdulkadiroğlu, A., Pathak, P., and Roth, E. (2009), "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 99, 1954-1978.


centralized matching context is novel and differentiates the current paper. There is a further difference. Jackson and Sonnenschein (2007) demonstrated the efficiency of linking when (linkable) decisions tend to infinity, relying largely on the logic of the law of large numbers. To our knowledge, the current paper is the first to characterize the precise welfare benefit of linking fixed (finite) number decisions (albeit with continuum of agents).

[18] In the finite economy, the agents have do not act as price takers. More generally, for a finite economy, there is no strategy-proof mechanism that treats the agents with the same preferences equally and implements an ex ante efficient allocation (Zhou, 1990).

[5] Abdulkadiroğlu, A., Pathak, P., Roth, E., and Sönmez, T. (2005), "The Boston Public School Match," *American Economic Review*, Papers and Proceedings, 95, 368-371.

[6] Abdulkadiroğlu, A., Pathak, P., Roth, E., and Sönmez, T. (2006), "Changing the Boston Mechanism: Strategyproofness as Equal Access," Unpublished manuscript, Harvard University.

[7] Abdulkadiroğlu, A., and Sönmez, T. (2003), "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729-747.

[8] Balinski, M. and Sönmez, T. (1999), "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84, 73-94.

[9] Bogomolnaia, A., and Moulin, H. (2001), "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295-328.

[10] Budish, E. and Cantillon, E. (2011), "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," Unpublished manuscript.

[11] Casella, A. (2005), "Storable Votes," *Games and Economic Behavior*, 51, 391-419.

[12] Che, Y.-K., and Gale, I. (2000), "The Optimal Mechanism for Selling to Budget-Constrained Buyer," *Journal of Economic Theory*, 92, 198-233.

[13] Che, Y.-K., and Kojima, F. (2010), "Asymptotic Equivalence of Random Priority and Probabilistic Serial Mechanisms," *Econometrica*, 78, 1625-1672.

[14] Dubins, L., and Freedman, D. (1981), "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 88, 485-494.

[15] Ehlers, L. (2006), "Respecting Priorities when Assigning Students to Schools," working paper, CIREQ.

[16] Erdil, A., and Ergin, H. (2008), "What's the Matter with Tie-breaking? Improving Efficiency with School Choice," *American Economic Review*, 98, 669-689.

[17] Ergin, H., and Sönmez, T. (2006), "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 90, 215-237.

[18] Gale, D., and Shapley, L. (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9-15.

[19] Hylland, A. and Zeckhauser, R. (1979), "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 87, 293-314.

[20] Jackson, M., and Sonnenschein, P. (2007), "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75, 241-257.

[21] Kesten, O. (2010), "School Choice with Consent," *Quarterly Journal of Economics*, 125, 1297-1348.

[22] Mas-Colell, A. (1984), "On a Theorem of Schmeidler," *Journal of Mathematical Economics*, 13, 201-206.

[23] Miralles, A. (2008), "School Choice: The Case for the Boston Mechanism," Unpublished manuscript.

[24] Pathak, P. and Sönmez, T. (2008), "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98, 1636-52.

[25] Roth, A. (1982), "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7, 617-628.

[26] Roth, A. (1985), "The College Admissions Problem is Not Equivalent to the Marriage Problem," *Journal of Economic Theory*, 36, 277-288.

[27] Roth, A. (2008), "Deferred Acceptance Algorithm: History, Theory, Practice, and Open Questions," *International Journal of Game Theory*, 36, 537-569.

[28] Sönmez, T., and Ünver, M.U. (forthcoming), "Course Bidding at Business Schools," *International Economic Review*.

[29] Zhou, L. (1990), "On a Conjecture by Gale about One-Sided Matching Problems," *Journal of Economic Theory*, 52, 123-135.

# Appendix: Proofs of the main results

**Proof of Lemma 3.** For any $S' \subset S$ and $s_a \in S'$, let $m_a(S') := \mu(\{\mathbf{v}|v_a \geq v_b, \forall s_b \in S'\})$ be the measure of students who prefer school $a$ the most among $S'$. The cutoffs of the schools are then defined recursively as follows. Let $\hat{S}^0 \equiv S$ $\hat{c}^0 \equiv 0$, and $\hat{x}_a^0 \equiv 0$ for every $a \in A$. Given $\hat{S}^0, \hat{c}^0, \{\hat{x}_a^0\}_{a\in A}, \ldots, \hat{S}^{t-1}, \hat{c}^{t-1}, \{\hat{x}_a^{t-1}\}_{a\in A}$, and for each $a \in A$ define

$$\hat{c}_a^t = \sup\left\{c \in [0,1] \middle| \hat{x}_a^{t-1} + m_a(\hat{S}^{t-1})\left(c - \hat{c}^{t-1}\right) < 1\right\}, \tag{3}$$

$$\hat{c}^t = \min_{s_i \in \hat{S}^{t-1}} \hat{c}_a^t, \tag{4}$$

$$\hat{S}^t = \hat{S}^{t-1} \setminus \{s_a \in \hat{S}^{t-1} | \hat{c}_a^t = c^t\}, \tag{5}$$

$$\hat{x}_a^t = \hat{x}_a^{t-1} + m_a(\hat{S}^{t-1})\left(\hat{c}^t - \hat{c}^{t-1}\right). \tag{6}$$

Each recursion step $t$ determines the cutoff of school(s) given cutoffs $\{\hat{c}^0, ..., \hat{c}^{t-1}\}$. Students with draw $\omega > \hat{c}^{t-1}$ can never be assigned to schools $S \setminus S^{t-1}$. For each school $a \in S^{t-1}$ with remaining capacity, a fraction $\hat{x}_a^{t-1}$ is claimed by students with draws less than $\hat{c}^{t-1}$, so only fraction $1 - \hat{x}_a^{t-1}$ of seats can be assigned to students with draws $\omega > \hat{c}^{t-1}$. If school $a$ has the next highest cutoff, $\hat{c}^t$, then the remaining capacity $1 - \hat{x}_a^{t-1}$ must equal the measure of those students who prefer $s_a$ the most among $S^{t-1}$ and have drawn numbers in $[\hat{c}^{t-1}, \hat{c}^t]$. This, together with the fact that school $a$ has cutoff $\hat{c}^t$, implies (3) and (4). The recursion definition implies (5) and (6).

The recursive equations uniquely determine the set of cutoffs $\{\hat{c}^0, ..., \hat{c}^k\}$, where $k \le n$. The cutoff for school $a \in S$ is then given by $c_a := \{\hat{c}^t | \hat{c}_a^t = \hat{c}^t\}$. It clearly follows from (3) and (4) for $t = 1$ that $\hat{c}^1 > 0$. It also easily follows that $\hat{c}^k = 1$. Obviously $\hat{c}^k \le 1$. We also cannot have $\hat{c}^k < 1$, or else there will be positive measure of students unassigned, which cannot occur since every student prefers each school to being unassigned, and the measure of all students coincides with the total capacity of schools.

Although it is possible for more than one school to have the same cutoff, this is not generic. If there are schools with the same cutoff, we must have $s_a \neq s_b \in \hat{S}^{t-1}$ for some $t$ and $\hat{S}^{t-1}$ such that $\hat{c}_a^t = \hat{c}_b^t$, which entails a loss of dimension for $\mathfrak{m}$ within $\mathfrak{M}$. Hence, the Lebesque measure of the set of $\mathfrak{m}$'s involving such a restriction is zero. It thus follows that generically no two schools have the same cutoff. ▌

**Proof of Lemma 4.** For each $a \in A$ and any $S' \subset S \setminus \{s_a\}$, let

$$m_a^{S'} := \mu(\{\mathbf{v} \in \mathcal{V} | v_b \ge v_a \ge v_c, \forall s_b \in S', \forall s_c \in S \setminus (S' \cup \{a\})\})$$

be the measure of those students whose preference order of school $a$ follows right after schools in $S'$. (Note that the order of schools within $S'$ does not matter here.) We can then define the conditions for cutoffs $\{\tilde{c}_1, ..., \tilde{c}_n\}$ under DA-MTB as the following system of simultaneous equations. Specifically, for any school $a \in S$, we must have

$$\tilde{c}_a\left(m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'}\left[\prod_{s_b \in S'}(1 - \tilde{c}_b)\right]\right) = 1. \tag{7}$$

The LHS has the measure of students admitted by school $a$. They consist of those students who prefer $s_a$ most and have admittable lottery draws for $s_a$ (i.e., $\omega_a \leq \tilde{c}_a$), and of those who prefer schools $S' \subset S \setminus \{s_a\}$ more than $s_a$ but have bad draws for those schools but have an admittable draw for school $a$. In equilibrium, the cutoffs must be such that these aggregate measures equal one (the capacity of school $a$).

To show that there exists a set $\{\tilde{c}_1, ..., \tilde{c}_n\}$ of cutoffs satisfying the system of equations (7), let $\Upsilon := (\Upsilon_1, ..., \Upsilon_n) : [0,1]^n \to [0,1]^n$ be a function whose $a$'s component is defined as:

$$\Upsilon_a(\tilde{c}_1, ..., \tilde{c}_n) = \min\left\{ \frac{1}{m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[\prod_{s_b \in S'} (1 - \tilde{c}_b)\right]}, 1 \right\},$$

where we adopt the convention that $\min\{\frac{1}{0}, 1\} = 1$.

Observe that self mapping $\Upsilon(\cdot)$ is a monotone increasing on a nonempty complete lattice. Hence, by the Tarski's fixed point theorem, there exists a largest fixed point $\mathbf{c}^* = (c_1^*, ..., c_n^*)$ such that $\Upsilon(\mathbf{c}^*) = \mathbf{c}^*$, and $\mathbf{c}^* \geq \tilde{\mathbf{c}}^*$ for any fixed point $\tilde{\mathbf{c}}^*$.

We now show that at any such fixed point $\tilde{\mathbf{c}}^*$,

$$\frac{1}{m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[\prod_{s_b \in S'} (1 - \tilde{c}_b^*)\right]} \leq 1, \tag{8}$$

for each $a \in A$. Suppose this is not the case for some $i$. Then, by the construction of the mapping, we must have $\tilde{c}_a^* = 1$. This means that all students are assigned to some schools. Therefore, by pure accounting,

$$\sum_{a \in A} \tilde{c}_a^* \left( m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[ \prod_{s_b \in S'} (1 - \tilde{c}_b^*) \right] \right) = n. \tag{9}$$

Yet, since (8) fails for some school,

$$\begin{aligned} &\sum_{a \in A} \tilde{c}_a^* \left( m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[ \prod_{s_b \in S'} (1 - \tilde{c}_b^*) \right] \right) \\ &< \sum_{a \in A} \left( \frac{1}{m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[\prod_{s_b \in S'} (1 - \tilde{c}_b^*)\right]} \right) \left( m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[ \prod_{s_b \in S'} (1 - \tilde{c}_b^*) \right] \right) = n, \end{aligned}$$

where the strict inequality follows since, for school $c$ for which (8) holds, $\tilde{c}_c^* = \frac{1}{m_c^{\emptyset} + \sum_{S' \subset S \setminus \{s_c\}} m_c^{S'} \left[\prod_{s_b \in S'} (1 - \tilde{c}_b^*)\right]}$ and, for school $a$ for which (8) does not hold, $\tilde{c}_a^* = 1 < \frac{1}{m_a^{\emptyset} + \sum_{S' \subset S \setminus \{s_a\}} m_a^{S'} \left[\prod_{s_b \in S'} (1 - \tilde{c}_b^*)\right]}$. This inequality contradicts (9). Since (8) holds for each $a \in A$, the fixed point $(\tilde{c}_1^*, ..., \tilde{c}_n^*)$ solves the system of equations (7). It is immediate from (7) that $\tilde{c}_a > 0, \forall a$. Further, there must exist a school $w \in S$ with $\tilde{c}_w = 1$, or else a positive measure of students are unassigned, which would violate (7). As before, it follows that the solutions to (7) are generically distinct.

To establish uniqueness, suppose to the contrary $\mathbf{c}^* > \tilde{\mathbf{c}}^*$: $c_b^* \geq \tilde{c}_b^*$ for all $b$ and $c_a^* > \tilde{c}_a^*$ for some $a$. Let $w \in S$ be such that $\tilde{c}_w^* = 1$. Since $\mathbf{c}^* \geq \tilde{\mathbf{c}}^*$, $c_w^* = 1$. Since (7) must be satisfied for $w$ under both cutoffs, we have

$$\left(m_w + \sum_{S' \subset S \setminus \{w\}} m_w^{S'} \left[\prod_{s_b \in S'} (1 - c_b)\right]\right) = \left(m_w + \sum_{S' \subset S \setminus \{s_a\}} m_w^{S'} \left[\prod_{s_b \in S'} (1 - \tilde{c}_b)\right]\right) = 1,$$

which holds if and only if $c_b = \tilde{c}_b$ for all $b$. ▌

**Proof of Theorem 2.** The proof is an application of Theorem 2 of Mas-Colell (1984). ▌

**Proof of Lemma 5.** Part (i) follows trivially since such a student can target that school and get assigned to it with probability one. To prove part (ii) consider any student of type $\mathbf{v}$, whose values are all distinct. There are $\mu$-a.e. such $\mathbf{v}$. Suppose her most-preferred school $\pi_1(\mathbf{v}) =: a$ is undersubscribed and not a worst school. It is then her best response to target $s_a$, since doing so can guarantee assignment to school $a$ for sure, whereas targeting some other school reduces her chance of assignment to that school. Hence, the student must be targeting $s_a$ in equilibrium.

To prove part (iii), consider any $\mathbf{v}$ (with distinct values), such that $\pi_1(\mathbf{v}) \neq w$. Suppose first $\sigma_a^*(\mathbf{v}) > 0$ for some oversubscribed school $a$. It follows from the above observation that she must strictly prefer school $a$ to all undersubscribed schools. Hence, she lists $s_a$ ahead of all undersubscribed schools in her ordinal list. Whenever she targets school $a$, she can never place in any oversubscribed school other than $s_a$, so she will apply to school $a$ with probability one. Suppose next $\sigma_b^*(\mathbf{v}) > 0$ for some undersubscribed school $b$. Then, the student must prefer $s_b$ to all other undersubscribed schools, so she will apply to school $b$ with probability one whenever she fails to place in any oversubscribed school she may list ahead of $s_b$ in the ordinal list. Whenever she targets school $b$, she is surely rejected by all oversubscribed schools she may list ahead of $s_b$, so she will apply to $s_b$ with probability one. We thus conclude that $\sigma^*(\mathbf{v}) = \bar{\sigma}^*(\mathbf{v})$ for $\mu$-a.e. $\mathbf{v}$. ▌

**Proof of Theorem 3:** To prove part (i), let school $b$ be such that $\tilde{c}_b = 1$. Hence, any students who prefer $s_b$ to $s_a$ can never be assigned to $s_a$. Hence, the allocation does not admit any trading cycle within $\{a, b\}$, and is thus OE within $\{a, b\}$ (Lemma 1). The allocation is then PE within $\{a, b\}$ by Lemma 2-(iii).

To prove part (ii), take any two schools $\{a, b\}$, with $\tilde{c}_a, \tilde{c}_b < 1$. There is a positive measure of students whose first- and second-most preferred schools are $a$ and $b$, respectively (call them "type-$a$"). Likewise, there is a positive measure of so-called "type-$b$" students whose first- and second-most preferred schools are $b$ and $a$, respectively. A positive measure of type-$a$ students draw $(\omega_a, \omega_b)$ such that $\omega_a > \tilde{c}_a$ and $\omega_b < \tilde{c}_b$; and a positive measure of type-$b$ students draw $(\omega'_a, \omega'_b)$ with $\omega'_a < \tilde{c}_a$ and $\omega'_b > \tilde{c}_b$. Clearly, the former type students are assigned to $b$ and the latter to $a$, so both types of students will benefit from swapping their assignments. Part (ii) then follows since generically there is only one school with cutoff equal to 1 (Lemma 4). ▌

**Proof of Theorem 4:** To prove part (i), suppose $s_a \rhd^{\phi^S} s_b$. Then, we must have $c_a < c_b$. Or else, any students who prefer school $a$ to $b$ can never be assigned to shool $b$. This is because any such student will rank $s_a$ ahead of $s_b$ (by strategyproofness), so if she is rejected by $s_a$, her draw must be $\omega > c_a \geq c_b$, not good enough for $s_b$. Hence, if $s_1 \rhd^{\phi^S} ... \rhd^{\phi^S} s_k \rhd^{\phi^S} s_1$, then $c_{s_1} < ... < c_{s_k} < c_{s_1}$, a contradiction. Hence, it is OE (and thus pairwise PE).

To prove part (ii), recall from Lemma 3 that the schools' cutoffs are generically distinct. Take any set $\{a, b, c\}$ with $c_a < c_b < c_c$. Then, by the full support assumption, there exists a positive measure of $\mathbf{v}$'s satisfying $v_a > v_b > v_c > v_d$ for all $d \neq a, b, c$. These students will then have a positive chance of being assigned to each school in $\{a, b, c\}$, for their draws will land in the intervals, $[0, c_a]$, $[c_a, c_b]$ and $[c_b, c_c]$, with positive probabilities. Again, given the full support assumption, such students will all differ in their marginal rate of substitution among the three schools. Then, just as with the motivating example, one can construct a mutually beneficial trading of shares of these schools among these students. ▌

**Proof of Theorem 5:** Part (i) builds on part (ii), so it will appear last. Throughout, we let $K$ and $J$ be the sets of over- and under-subscribed schools.

*Part (ii):* Let $\sigma^*(\cdot)$ be an equilibrium and $\phi^*(\cdot)$ be the associated allocation. For any $\mathbf{v} \in \mathcal{V}$, consider an optimization problem:

$$[P(\mathbf{v})] \qquad \max_{\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}} \sum_{a \in A} v_a x_a \ \text{ subject to } \sum_{s_a \in K} p_a x_a \leq \sum_{s_a \in K} p_a \phi^*_a(\mathbf{v}),$$

where $p_a \equiv \max\{\int \overline{\sigma}^*_a(\tilde{\mathbf{v}}) d\mu(\tilde{\mathbf{v}}), 1\}$.

We first prove that $\phi^*(\mathbf{v})$ solves $[P(\mathbf{v})]$. This is trivially true for any type $\mathbf{v}$-student whose most preferred school is the worst school $w$. Then, by Lemma 5-(i), $\phi^*_w(\mathbf{v}) = 1$ and $x_a = \phi^*_a(\mathbf{v}) = 0, \forall s_a \in K$. So, $\phi^*(\mathbf{v})$ solves $[P(\mathbf{v})]$.

Hence, assume that $\pi_1(\mathbf{v}) \neq w$ in what follows. Fix any such $\mathbf{v}$, and fix any arbitrary $\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}$ satisfying the constraint of $[P(\mathbf{v})]$. We show below that the type $\mathbf{v}$ student can mimic the assignment $\mathbf{x}$ by adopting a certain targeting strategy in the CADA game, assuming that all other players play their equilibrium strategies $\sigma^*$.

To begin, consider a strategy called $s_a$ in which she targets school $a \in S$ and also top-ranks it in her ordinal list but ranks all other schools truthfully. If type $\mathbf{v}$ plays strategy $s_a$, then she will be assigned to school $a$ with probability

$$\frac{1}{\max\{\int \overline{\sigma}^*_a(\tilde{\mathbf{v}}) d\mu(\tilde{\mathbf{v}}), 1\}} = \frac{1}{p_a}.$$

If $s_a \in J$, this probability is one. If $s_a \in K$, then she will be rejected by school $a$ with positive probability. If she is rejected, she will apply to other schools. Clearly, she will not succeed in getting into any schools in $K$, since they are oversubscribed. The conditional probabilities of getting assigned to schools $J$ do not depend on which school in $K$ she has targeted (and gotten turned down), due to our design whereby her non-target draw $\omega_R$ is independent of her target draw $\omega_T$ (recall footnote 10). For each $s_b$, let that conditional assignment probability be $\bar{\phi}_b^*(\mathbf{v})$ for type $\mathbf{v}$. Obviously, $\sum_{s_b \in J} \bar{\phi}_b^*(\mathbf{v}) = 1$.

Suppose the type $\mathbf{v}$ student randomizes by choosing "strategy $s_a$" with probability $y_a := p_a x_a$, for each $s_a \in K$, and with probability

$$y_b := \sigma_b^*(\mathbf{v}) + \left[\sum_{s_a \in K} (\sigma_a^*(\mathbf{v}) - p_a x_a)\left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v}),$$

for each $s_b \in J$. Observe $y_b \geq 0$ for all $s_b \in S$. This is obvious for $s_b \in K$. For $s_b \in J$, this follows since the terms in the square brackets are nonnegative:

$$\begin{aligned}
\sum_{s_a \in K} (\sigma_a^*(\mathbf{v}) - p_a x_a)\left(1 - \frac{1}{p_a}\right) &= \sum_{s_a \in K} (p_a \phi_a^*(\mathbf{v}) - p_a x_a)\left(1 - \frac{1}{p_a}\right) \\
&= \left[\sum_{s_a \in K} p_a(\phi_a^*(\mathbf{v}) - x_a)\right] - \left[\sum_{s_a \in K} (\phi_a^*(\mathbf{v}) - x_a)\right] = \sum_{s_a \in K} p_a(\phi_a^*(\mathbf{v}) - x_a) \geq 0,
\end{aligned}$$

where the first equality is implied by Lemma 5-(iii), the third equality holds since $\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}$ (which implies $\sum_{s_a \in K} x_a = \sum_{s_a \in K} \phi^*(\mathbf{v})$), and the last inequality follows from the fact that $\mathbf{x}$ satisfies the constraint of $[P(\mathbf{v})]$. Further,

$$\begin{aligned}
\sum_{a \in A} y_a &= \sum_{s_a \in K} p_a x_a + \sum_{s_b \in J} \left[\sigma_b^*(\mathbf{v}) + \left[\sum_{s_a \in K} (\sigma_a^*(\mathbf{v}) - p_a x_a)\left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v})\right] \\
&= \sum_{s_a \in K} p_a x_a + \sum_{s_b \in J} \sigma_b^*(\mathbf{v}) + \sum_{s_a \in K} \left[(\sigma_a^*(\mathbf{v}) - p_a x_a)\left(1 - \frac{1}{p_a}\right)\right] \left(\sum_{s_b \in J} \bar{\phi}_b^*(\mathbf{v})\right) \\
&= \sum_{s_a \in K} p_a x_a + \sum_{s_b \in J} \sigma_b^*(\mathbf{v}) + \sum_{s_a \in K} \left[(\sigma_a^*(\mathbf{v}) - p_a x_a)\left(1 - \frac{1}{p_a}\right)\right] \\
&= \sum_{i \in K} \sigma_i^*(\mathbf{v}) + \sum_{j \in J} \sigma_j^*(\mathbf{v}) + \sum_{i \in K} (\phi_i^*(\mathbf{v}) - x_i) = \sum_{i \in S} \sigma_i^*(\mathbf{v}) = 1.
\end{aligned}$$

The third equality holds since $\sum_{s_b \in J} \bar{\phi}_b^*(\mathbf{v}) = 1$, the fourth is implied by Lemma 5-(iii), and the fifth follows since $\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}$.

By playing the mixed strategy $(y_1, ..., y_n)$, the student is assigned to school $a \in K$ with probability

$$\frac{y_a}{p_a} = x_a,$$

and to each school $b \in J$ with probability

$$
\begin{aligned}
& y_b + \left[\sum_{s_a \in K} y_a \left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v}) \\
&= \sigma_b^*(\mathbf{v}) + \left[\sum_{s_a \in K} (\sigma_a^*(\mathbf{v}) - p_a x_a) \left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v}) + \left[\sum_{s_a \in K} p_a x_a \left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v}) \\
&= \sigma_b^*(\mathbf{v}) + \left[\sum_{s_a \in K} \sigma_a^*(\mathbf{v}) \left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v}) = \overline{\sigma}_b^*(\mathbf{v}) + \left[\sum_{s_a \in K} \overline{\sigma}_a^*(\mathbf{v}) \left(1 - \frac{1}{p_a}\right)\right] \bar{\phi}_b^*(\mathbf{v}) \\
&= \phi_j^*(\mathbf{v}) = x_j.
\end{aligned}
$$

In other words, the type $\mathbf{v}$ student can mimic any $\mathbf{x} \in \Delta^K_{\phi^*(\mathbf{v})}$ that satisfies $\sum_{s_a \in K} p_a x_a \leq \sum_{s_a \in K} p_a \phi_a^*(\mathbf{v})$ by playing a certain strategy available in the CADA game. Since every feasible $\mathbf{x}$ can be mimicked by a strategy available in the equilibrium of CADA, $\phi^*(\cdot)$ is a best response for type $\mathbf{v}$, and since it and satisfies the constraints of $[P(\mathbf{v})]$, $\phi^*(\cdot)$ must solve $[P(\mathbf{v})]$.

Moreover, since $\mu$ is atomless and $[P(\mathbf{v})]$ has a linear objective function on a convex set, $\phi^*(\mathbf{v})$ must be the unique solution to $[P(\mathbf{v})]$ for a.e. $\mathbf{v}$.

We prove the statement of the theorem by contradiction. Suppose to the contrary that there exists an allocation $\phi(\cdot) \in \mathcal{X}^K_{\phi^*}$ that Pareto dominates $\phi^*(\cdot)$. Then, for a.e. $\mathbf{v}$, $\phi(\mathbf{v})$ must either solve $[P(\mathbf{v})]$ or violate its constraints. For a.e. $\mathbf{v}$, the solution to $[P(\mathbf{v})]$ is unique and coincides with $\phi^*(\mathbf{v})$. This implies that for a.e. $\mathbf{v}$,

$$\sum_{s_a \in K} p_a \phi_a(\mathbf{v}) \geq \sum_{s_a \in K} p_a \phi_a^*(\mathbf{v}). \tag{10}$$

Further, for $\phi$ to Pareto-dominate $\phi^*$, there must exist a set $A \subset \mathcal{V}$ with $\mu(A) > 0$ such that each student $\mathbf{v} \in A$ must strictly prefer $\phi(\mathbf{v})$ to $\phi^*(\mathbf{v})$, which must imply (since $\phi^*(\mathbf{v})$ solves $[P(\mathbf{v})]$)

$$\sum_{s_a \in K} p_a \phi_a(\mathbf{v}) > \sum_{s_a \in K} p_a \phi_a^*(\mathbf{v}), \, \forall \mathbf{v} \in A. \tag{11}$$

Combining (10) and (11), we get

$$\sum_{s_a \in K} p_a \int \phi_a(\mathbf{v}) d\mu(\mathbf{v}) > \sum_{s_a \in K} p_a \int \phi_a^*(\mathbf{v}) d\mu(\mathbf{v}). \tag{12}$$

Now since $\phi(\cdot) \in \mathcal{X}$, for each $a \in A$,

$$\int \phi_a(\mathbf{v}) d\mu(\mathbf{v}) = 1 = \int \phi_a^*(\mathbf{v}) d\mu(\mathbf{v}).$$

Multiplying both sides by $p_a$ and summing over $K$, we get

$$\sum_{s_a \in K} p_a \int \phi_a(\mathbf{v}) d\mu(\mathbf{v}) = \sum_{s_a \in K} p_a \int \phi_a^*(\mathbf{v}) d\mu(\mathbf{v}),$$

which contradicts (12). We thus conclude that $\phi^*$ is Pareto optimal within $K$.

*Part (iii):* Consider the following maximization problem for every $\mathbf{v} \in \mathcal{V}$:

$$[\overline{P}(\mathbf{v})] \qquad \max_{\mathbf{x}\in\Delta} \sum_{a\in A} v_a x_a \;\text{ subject to }\; \sum_{s_a\in K} p_a x_a \leq 1.$$

When we have only one undersubscribed school, say $b$, then its assignment is determined by $x_b = 1 - \sum_{s_a\in K} x_a$. Therefore, an assignment $\mathbf{x} \in \Delta$ is feasible in CADA game if (and only if) the constraint of $[\overline{P}(\mathbf{v})]$ holds.

Now consider the following maximization problem:

$$[\overline{P}'(\mathbf{v})] \qquad \max_{\mathbf{x}\in\Delta} \sum_{a\in A} v_a x_a \;\text{ subject to }\; \sum_{s_a\in K} p_a x_a \leq \sum_{s_a\in K} p_a \phi_a^*(\mathbf{v}).$$

Since $\phi^*(\cdot)$ solves a less constrained problem $[\overline{P}(\mathbf{v})]$ and is still feasible in $[\overline{P}'(\mathbf{v})]$, it must be an optimal solution for $[\overline{P}'(\mathbf{v})]$. The rest of the proof is shown by the same argument as in Part (ii).

*Part (i):* The argument in the text already established that the allocation cannot admit a trading cycle that includes both oversubscribed and unsubscribed schools. It cannot admit a trading cycle comprising only oversubscribed schools, since the allocation is PE within these schools, by Part (ii), making it OE within the schools, by Lemma 2-(ii). It cannot admit a trading cycle comprising only undersubscribed schools, since the logic of Theorem 4-(i) implies that it is OE within undersubscribed schools. Since the allocation cannot admit any trading cycle, it must be OE. ▌

**Proof of Theorem 6:** Consider first a DA algorithm with any random tie-breaking. Since all students submit the same ranking of the schools, they are assigned to each school with the same probability $1/n$. In other words, the allocation is $\phi^{DA}(\mathbf{v}) = (\frac{1}{n}, ..., \frac{1}{n})$ for all $\mathbf{v}$.

Consider now CADA algorithm and an associated equilibrium $\sigma^*$. Then, a fraction $\alpha_a^* := \int \sigma_a^*(\mathbf{v}) d\mu(\mathbf{v})$ of students target $s_a \in S$ in equilibrium. The equilibrium induces a mapping $\varphi^* : S \mapsto \Delta$, such that a student is assigned to school $b$ with probability $\varphi_b^*(a)$ if she targets $s_a$.

Since the capacity of each school is filled in equilibrium, we must have, for each $s_b \in S$,

$$\sum_{a\in A} \alpha_a^* \varphi_b^*(a) = 1. \tag{13}$$

That is, a measure $\alpha_i^*$ of students target $s_a$, and a fraction $\varphi_b^*(a)$ of those is assigned to school $b$. Summing the product over all $s_a$ then gives the measure of students assigned to $s_b$, which must equal its capacity, 1.

Consider a student with any arbitrary $\mathbf{v} \in \mathcal{V}$. We show that there is a strategy she can employ to mimic the random assignment $\phi^{DA}$. Suppose she randomizes by targeting school $a$

with probability

$$y_a := \frac{\alpha_a^*}{\sum_b \alpha_b^*} = \frac{\alpha_a^*}{n}.$$

Then, the probability that she will be assigned to any school $k$ is

$$\sum_b y_b \varphi_k^*(b) = \sum_b \frac{\alpha_b^*}{n} \varphi_k^*(b) = \frac{1}{n},$$

where the second equality follows from (13). That is, she can replicate the same ex ante assignment with the randomization strategy as $\phi^{DA}(\mathbf{v})$. Hence, the student must be at least weakly better off under CADA. ▌

**Figure 3: Welfare as Percentage of First Best**

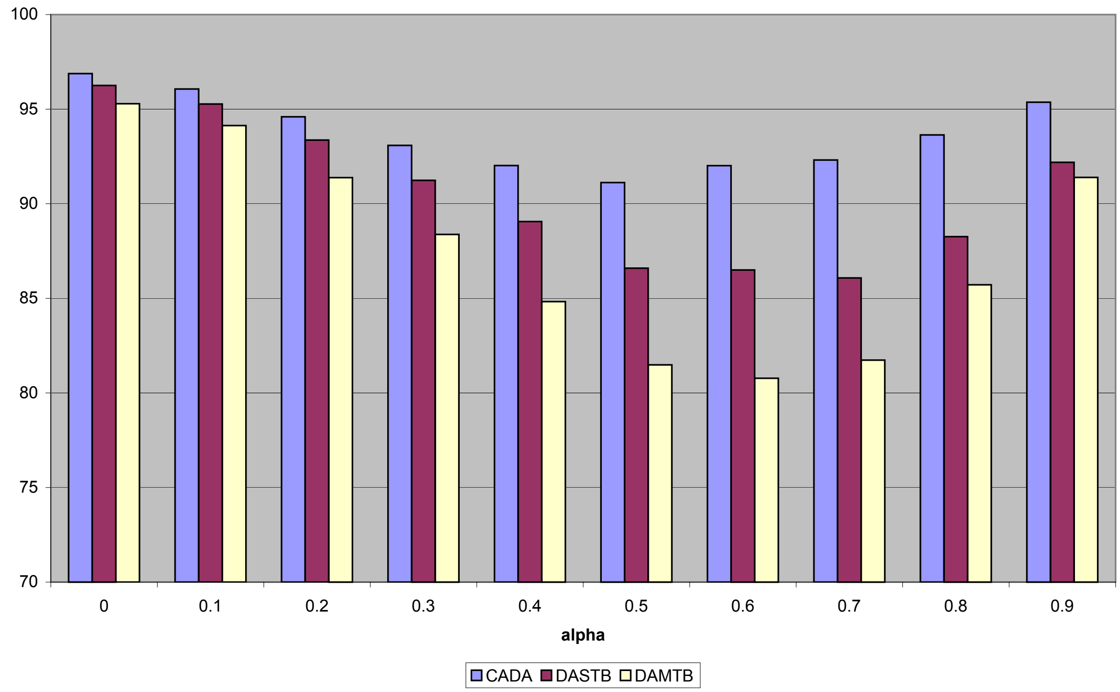


**Figure 4: Percentage of Students Getting Their First Choice**

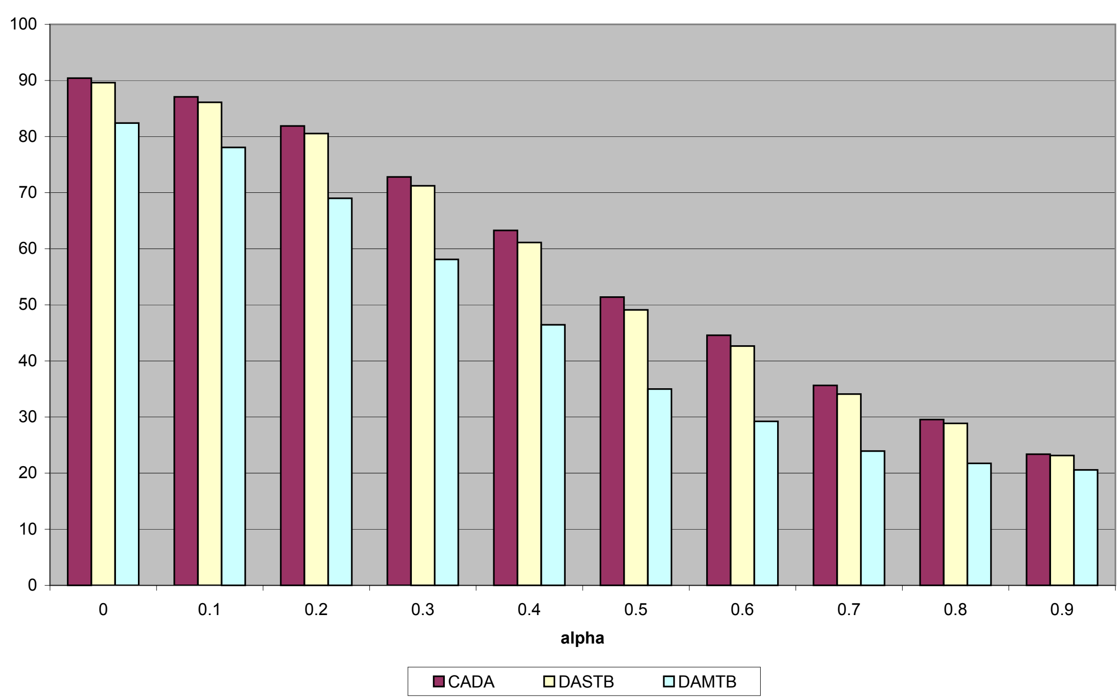

Figure 5: Average Utility of Receivers of kth Choice, CADA vs DASTB

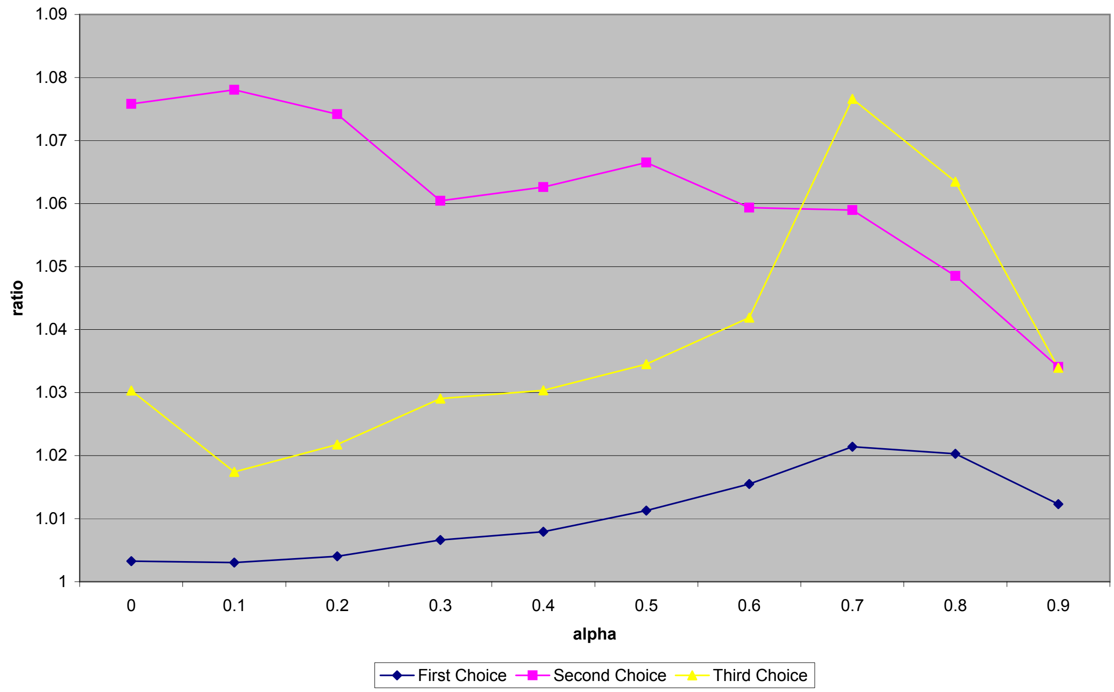


Figure 6: Average Number of Popular Schools and Oversubsribed Schools

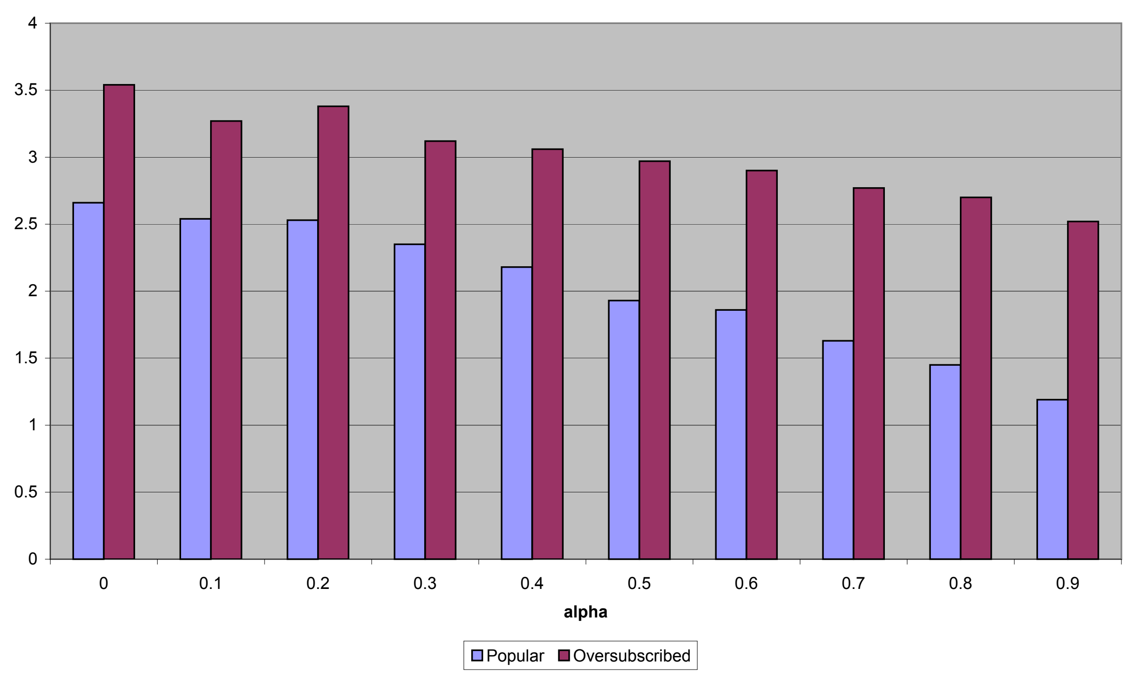

**Figure 7: Welfare as Percentage of First Best - with priorities**