Appendix

A Main Proofs

A.1 Proof of Theorem 1

Let $F_v^i(r)$ denote the cumulative distribution function (CDF) of R_{iv} evaluated at r and define

$$F_v(r|\theta) = E[F_v^i(r)|\theta_i = \theta]. \tag{17}$$

This is the fraction of type θ applicants with tie-breaker v below r (set to zero when type θ ranks no schools using tie-breaker v). We may condition on additional events.

Recall that the joint distribution of tie-breakers for applicant i is assumed to be continuously differentiable with positive density. This assumption has the following implication: The conditional distribution of tie-breaker v, $F_v(r|e)$, is continuously differentiable, with $F_v'(r|e) > 0$ at any $r = \tau_1, ..., \tau_S$. Here, the conditioning event e is any event of the form that $\theta_i = \theta, R_{iu} > r_u$ for $u = 1, ..., v-1$, and $T_i(\delta) = T$.

Take any large market with the general tie-breaking structure in Section 4. For each $\delta > 0$ and each tie-breaker $v = U+1, ..., V+1$, let $e(v)$ be short-hand notation for "$\theta_i = \theta, R_{iu} > MID_{\theta s}^u$ for $u = 1, ..., v-1, T_i(\delta) = T$, and $W_i = w$." Similarly, $e(1)$ is short-hand notation for "$\theta_i = \theta, T_i(\delta) = T$, and $W_i = w$." Let $\psi_s(\theta, T, \delta, w) \equiv E[D_i(s)|e(1)]$ be the assignment probability for an applicant with $\theta_i = \theta, T_i(\delta) = T$, and characteristics $W_i = w$. Our proofs use a lemma that describes this assignment probability. To state the lemma, for $v > U$, let

$$\Phi_\delta(v) \equiv \begin{cases} \dfrac{F_v(MID_{\theta s}^v + \delta|e(v)) - F_v(MID_{\theta s}^v - \delta|e(v))}{F_v(MID_{\theta s}^v|e(v)) - F_v(MID_{\theta s}^v - \delta|e(v))} & \text{if } t_b(\delta) = c \text{ for some } b \in B_{\theta s}^v \\ 1 & \text{otherwise.} \end{cases}$$

We use this object to define $\Phi_\delta \equiv \prod_{v=1}^{U}(1 - MID_{\theta s}^v) \prod_{v=U+1}^{V} \Phi_\delta(v)$. Finally, let

$$\Phi_\delta' \equiv \begin{cases} \max\left\{0, \dfrac{F_{v(s)}(\tau_s|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}{F_{v(s)}(\tau_s + \delta|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}\right\} & \text{if } v(s) > U \\ \max\left\{0, \dfrac{\tau_s - MID_{\theta s}^{v(s)}}{1 - MID_{\theta s}^{v(s)}}\right\} & \text{if } v(s) \leq U. \end{cases}$$

Lemma 1. *In the general tie-breaking setting of Section 4, for any fixed $\delta > 0$ such that*

$\delta < \min_{\theta,s,v} |\tau_s - MID^v_{\theta,s}|$, *we have:*

$$\psi_s(\theta, T, \delta, w) = \begin{cases} 0 & \text{if } t_s(\delta) = n \text{ or } t_b(\delta) = a \text{ for some } b \in B_{\theta s}, \\ \Phi_\delta & \text{otherwise and } t_s(\delta) = a, \\ \Phi_\delta \times \Phi'_\delta & \text{otherwise and } t_s(\delta) = c. \end{cases}$$

Proof of Lemma 1. We start verifying the first line in $\psi_s(\theta, T, \delta, w)$. Applicants who don't rank s have $\psi_s(\theta, T, \delta, w) = 0$. Among those who rank s, those of $t_s(\delta) = n$ have $\rho_{\theta s} > \rho_s$ or, if $v(s) \neq 0$, $\rho_{\theta s} = \rho_s$ and $R_{iv(s)} > \tau_s + \delta$. If $\rho_{\theta s} > \rho_s$, then $\psi_s(\theta, T, \delta, w) = 0$. Even if $\rho_{\theta s} \leq \rho_s$, as long as $\rho_{\theta s} = \rho_s$ and $R_{iv(s)} > \tau_s + \delta$, student i never clears the cutoff at school s so $\psi_s(\theta, T, \delta, w) = 0$.

To show the remaining cases, take as given that it is not the case that $t_s(\delta) = n$ or $t_b(\delta) = a$ for some $b \in B_{\theta s}$. Applicants with $t_b(\delta) \neq a$ for all $b \in B_{\theta s}$ and $t_s(\delta) = a$ or c may be assigned $b \in B_{\theta s}$, where $\rho_{\theta b} = \rho_b$. Since the (aggregate) distribution of tie-breaking variables for type θ students is $\hat{F}_v(\cdot|\theta) = F_v(\cdot|\theta)$, conditional on $T_i(\delta) = T$, the proportion of type θ applicants not assigned any $b \in B_{\theta s}$ where $\rho_{\theta b} = \rho_b$ is $\Phi_\delta = \prod_{v=1}^{U}(1 - MID^v_{\theta s}) \prod_{v=U+1}^{V} \Phi_\delta(v)$ since each $\Phi_\delta(v)$ is the probability of not being assigned to any $b \in B^v_{\theta s}$. To see why $\Phi_\delta(v)$ is the probability of not being assigned to any $b \in B^v_{\theta s}$, note that if $t_b(\delta) \neq c$ for all $b \in B^v_{\theta s}$, then $t_b(\delta) = n$ for all $b \in B^v_{\theta s}$ so that applicants are never assigned to any $b \in B^v_{\theta s}$. Otherwise, i.e., if $t_b(\delta) = c$ for some $b \in B^v_{\theta s}$, then applicants are assigned to s if and only if their values of tie-breaker v clear the cutoff of the school that produces $MID^v_{\theta s}$, where applicants have $t_s(\delta) = c$. This event happens with probability

$$\frac{F_v(MID^v_{\theta s}|e(v)) - F_v(MID^v_{\theta s} - \delta|e(v))}{F_v(MID^v_{\theta s} + \delta|e(v)) - F_v(MID^v_{\theta s} - \delta|e(v))},$$

implying that $\Phi_\delta(v)$ is the probability of not being assigned to any $b \in B^v_{\theta s}$.

Given this fact, to see the second line, note that every applicant of type $t_s(\delta) = a$ who is not assigned a higher choice is assigned s for sure because $\rho_{\theta s} < \rho_s$ or $\rho_{\theta s} + R_{iv(s)} < \xi_s$. Therefore, we have

$$\psi_s(\theta, T, \delta, w) = \Phi_\delta.$$

Finally, consider applicants with $t_s(\delta) = c$. The fraction of those who are not assigned a higher choice is Φ_δ, as explained above. Also, for tie-breaker $v(s)$, the tie-breaker values of these applicants are larger (worse) than $MID^{v(s)}_{\theta s}$. If $\tau_s < MID^{v(s)}_{\theta s}$, then no such applicant is assigned s. If $\tau_s \geq MID^{v(s)}_{\theta s}$, then the fraction of applicants who are assigned s conditional on $\tau_s \geq MID_{\theta s}^{v(s)}$ is given by

$$\max\left\{0, \frac{F_{v(s)}(\tau_s|e(V+1)) - \max\{F_{v(s)}(MID_{\theta s}^{v(s)}|e(V+1)), F_{v(s)}(\tau_s - \delta|e(V+1))\}}{F_{v(s)}(\tau_s+\delta|e(V+1)) - \max\{F_{v(s)}(MID_{\theta s}^{v(s)}|e(V+1)), F_{v(s)}(\tau_s - \delta|e(V+1))\}}\right\} \text{ if } v(s) > U$$

and

$$\max\left\{0, \frac{\tau_s - MID_{\theta s}^{v(s)}}{1 - MID_{\theta s}^{v(s)}}\right\} \text{ if } v(s) \leq U.$$

If $MID_{\theta s}^{v(s)} < \tau_s$, then $\delta < \min_{\theta,s,v} |\tau_s - MID_{\theta,s}^v|$ implies $MID_{\theta s}^{v(s)} < \tau_s - \delta$. This in turn implies

$$\max\{F_{v(s)}(MID_{\theta s}^{v(s)}|e(V+1)), F_{v(s)}(\tau_s - \delta|e(V+1))\} = F_{v(s)}(\tau_s - \delta|e(V+1)).$$

If $MID_{\theta s}^{v(s)} > \tau_s$, then $\delta < \min_{\theta,s,v} |\tau_s - MID_{\theta,s}^v|$ implies $MID_{\theta s}^{v(s)} > \tau_s + \delta$. By the definition of $e(V+1)$, $R_{iu} > MID_{\theta s}^u$ for $u = 1, ..., V$. Therefore, there is no applicant with $R_{iv(s)} > MID_{\theta s}^{v(s)}$ and $R_{iv(s)} \in [\tau_s - \delta, \tau_s + \delta]$.

Hence, conditional on $t_s(\delta) = c$ and not being assigned a choice preferred to s, the probability of being assigned s is given by Φ'_δ. Therefore, for students with $t_s(\delta) = c$, we have $\psi_s(\theta, T, \delta, w) = \Phi_\delta \times \Phi'_\delta$. □

Lemma 2. *In the general tie-breaking setting of Section 4, for all s, θ, and sufficiently small $\delta > 0$, we have:*

$$\psi_s(\theta, T, \delta, w) = \begin{cases} 0 & \text{if } t_s(0) = n \text{ or } t_b(0) = a \text{ for some } b \in B_{\theta s}, \\ \Phi_\delta^* & \text{otherwise and } t_s(0) = a, \\ \Phi_\delta^* \times & \dfrac{F_{v(s)}(\tau_s|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}{F_{v(s)}(\tau_s + \delta|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))} \\ & \text{otherwise and } t_s(0) = c \text{ and } v(s) > U. \\ \Phi_\delta^* \times & \max\{0, \dfrac{\tau_s - MID_{\theta s}^{v(s)}}{1 - MID_{\theta s}^{v(s)}}\} \\ & \text{otherwise and } t_s(0) = c \text{ and } v(s) \leq U. \end{cases} \tag{18}$$

where

$$\Phi_\delta^*(v) \equiv \begin{cases} \dfrac{F_v(MID_{\theta s}^v + \delta|e(v)) - F_v(MID_{\theta s}^v|e(v))}{F_v(MID_{\theta s}^v + \delta|e(v)) - F_v(MID_{\theta s}^v - \delta|e(v))} & \text{if } MID_{\theta s}^v = \tau_b \text{ and } t_b = c \text{ for some } b \in B_{\theta s}^v, \\ 1 & \text{otherwise} \end{cases}$$

and

$$\Phi^*_\delta \equiv \prod_{v=1}^{U}(1 - MID^v_{\theta s}) \prod_{v=U+1}^{V} \Phi^*_\delta(v).$$

Proof of Lemma 2. The first line follows from Lemma 1 and the fact that $t_s(0) = n$ or $t_b(0) = a$ for some $b \in B_{\theta s}$ imply $t_s(\delta) = n$ or $t_b(\delta) = a$ for some $b \in B_{\theta s}$ for sufficiently small $\delta > 0$.

For the remaining lines, first note that conditional on $t_s(0) \neq n$ and $t_b(0) \neq a$ for all $b \in B_{\theta s}$, we have $\Phi^*_\delta(v) = \Phi_\delta(v)$ and so $\Phi^*_\delta = \Phi_\delta$ holds for small enough δ. Φ^*_δ therefore is the probability of not being assigned to a school preferred to s in the last three cases.

The second line is then by the fact that $t_s(0) = a$ implies $t_s(\delta) = a$ for small enough $\delta > 0$. The third line is by the fact that for small enough $\delta > 0$,

$$\begin{aligned}\Phi'_\delta &= \max\left\{0, \frac{F_{v(s)}(\tau_s|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}{F_{v(s)}(\tau_s + \delta|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}\right\} \\ &= \frac{F_{v(s)}(\tau_s|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}{F_{v(s)}(\tau_s + \delta|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))},\end{aligned}$$

where we invoke Assumption 2, which implies $MID^v_{\theta s} \neq \tau_s$. The last line directly follows from Lemma 1. □

We use Lemma 2 to derive Theorem 1. We characterize $\lim_{\delta\to 0}\psi_s(\theta, T, \delta, w)$ and show that it coincides with $\psi_s(\theta, T)$ in the main text. In the first case in Lemma 2, $\psi_s(\theta, T, \delta, w)$ is constant (0) for any small enough δ. The constant value is also $\lim_{\delta\to 0}\psi_s(\theta, T, \delta, w)$ in this case.

To characterize $\lim_{\delta\to 0}\psi_s(\theta, T, \delta, w)$ in the remaining cases, note that by the differentiability of $F_v(\cdot|e(v))$ (recall the continuous differentiability of $F^i_v(r|e)$), L'Hopital's rule implies:

$$\lim_{\delta\to 0}\frac{F_{v(s)}(\tau_s|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))}{F_{v(s)}(\tau_s + \delta|e(V+1)) - F_{v(s)}(\tau_s - \delta|e(V+1))} = \frac{F'_{v(s)}(\tau_s|e(V+1))}{2F'_{v(s)}(\tau_s|e(V+1))} = 0.5$$

and

$$\lim_{\delta\to 0}\frac{F_v(MID^v_{\theta s} + \delta|e(v)) - F_v(MID^v_{\theta s}|e(v))}{F_v(MID^v_{\theta s} + \delta|e(v)) - F_v(MID^v_{\theta s} - \delta|e(v))} = \frac{F'_v(MID^v_{\theta s}|e(v))}{2F'_v(MID^v_{\theta s}|e(v))} = 0.5.$$

This implies $\lim_{\delta\to 0}\Phi^*_\delta(v) = 0.5^{1\{MID^v_{\theta s}=\tau_b \text{ and } t_b=c \text{ for some } b\in B^v_{\theta s}\}}$ since $1\{MID^v_{\theta s} = \tau_b$ and $t_b = c$ for some $b \in B^v_{\theta s}\}$ does not depend on δ. Therefore

$$\lim_{\delta\to 0}\Phi^*_\delta = \prod_{v=1}^{U}(1 - MID^v_{\theta s})0.5^{m_s(\theta,T)}$$

where $m_s(\theta, T) = |\{v > U : MID^v_{\theta s} = \tau_b \text{ and } t_b = c \text{ for some } b \in B^v_{\theta s}\}|$.

Combining these limiting facts with the fact that the limit of a product of functions equals the product of the limits of the functions, we obtain the following: $\lim_{\delta\to 0}\psi_s(\theta, T, \delta, w) = 0$ if (a) $t_s = n$ or (b) $t_b = a$ for some $b \in B_{\theta s}$. Otherwise,

$$\psi_s(\theta, T) = \begin{cases} \sigma_s(\theta, T)\lambda_s(\theta) & \text{if } t_s = a \\ \sigma_s(\theta, T)\lambda_s(\theta)\max\left\{0, \frac{\tau_s - MID_{\theta s}^{v(s)}}{1 - MID_{\theta s}^{v(s)}}\right\} & \text{if } t_s = c \text{ and } v(s) \leq U \\ 0.5\sigma_s(\theta, T)\lambda_s(\theta) & \text{if } t_s = c \text{ and } v(s) > U. \end{cases}$$

This expression coincides with $\psi_s(\theta, T)$, completing the proof of Theorem 1.

A.2 Proof of Corollary 1

Theorem 1 implies the following limiting conditional independence property:

$$\lim_{\delta\to 0} E[D_{Ai}|\psi_A(\theta, T, \delta), W_i] = \lim_{\delta\to 0} E[D_{Ai}|\psi_A(\theta, T, \delta)],$$

while the Corollary presumes exclusion, that is, we assume this holds for $W_i = Y_{0i}$. By the symmetry of conditional independence (Dawid, 1979), this implies

$$\lim_{\delta\to 0}(E[Y_i|D_{Ai} = 1, \psi_A(\theta, T, \delta) = p] - E[Y_i|D_{Ai} = 0, \psi_A(\theta, T, \delta) = p])$$

$$= \beta \lim_{\delta\to 0}(E[C_i|D_{Ai} = 1, \psi_A(\theta, T, \delta) = p] - E[C_i|D_{Ai} = 0, \psi_A(\theta, T, \delta) = p]),$$

where p is any value in $(0, 1)$ such that the first-stage effect $\lim_{\delta\to 0} E[C_i|D_{Ai} = 1, \psi_A(\theta, T, \delta) = p] - E[C_i|D_{Ai} = 0, \psi_A(\theta, T, \delta) = p] \neq 0$. Since we assume the first-stage effect is non-zero, the conclusion follows.

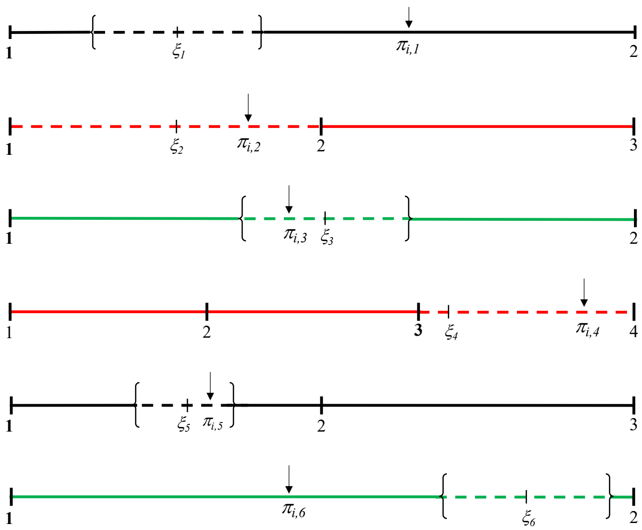
Online Appendices

B Understanding Theorem 1

Figure B1 illustrates Theorem 1 for an applicant who ranks screened schools 1, 3, 5 and 6 and lottery schools 2 and 4, where school k is applicant's k-th choice. The line next to each school represents applicant position (priority plus tie-breaker) for each school. Schools with the same colored lines have the same tie-breaker. Schools 1 and 5 use screened tie-breaker 2. Schools 2 and 4 use lottery tie-breaker 1. Schools 3 and 6 use screened tie-breaker 3.

Since school 1 has only one priority, positions run from 1 to 2. School *2* has two priority groups, so positions run from 1 to 3. Figure B1 indicates the applicants position π by an arrow. At screened schools, the brackets around the DA cutoff ξ represent the δ-neighborhood around the cutoff.

Figure B1: Illustrating Theorem 1



School	Position (priority + tie-breaker)	Tie-breaker	$v(s)$	$t_{is}(\delta)$	$\tau_s = \xi_s - \rho_s$	$\lambda_s(\theta)$	Schools contributing to $\lambda_s(\theta)$	$\sigma_s(\theta, T)$	Schools contributing to $\sigma_s(\theta, T)$	Assignment probability
1		Screened	2	*n*	ξ_1 - 1		n/a		n/a	0
2		Lottery	1	*c*	ξ_2 - 1	1	∅	1	∅	τ_2
3		Screened	3	*c*	ξ_3 - 1	1 - τ_2	*2*	1	∅	0.5(1 - τ_2)
4		Lottery	1	*c*	ξ_4 - 3	1 - τ_2	*2*	0.5	*3*	0.5max{0, τ_4 - τ_2 }
5		Screened	2	*c*	ξ_5 - 1	1 - max{ τ_2, τ_4}	*2, 4*	0.5	*3*	$(0.5)^2$ (1 - max{ τ_2, τ_4})
6		Screened	3	*a*	ξ_6 - 1	1 - max{ τ_2, τ_4}	*2, 4*	0.5^2	*3, 5*	$(0.5)^2$ (1 - max{ τ_2, τ_4})

Notes: This figure illustrates Theorem 1 for one applicant listing six schools. The applicant has marginal priority (shown in bold) at each. Dashes mark intervals in which offer risk is strictly between 0 and 1. The set of applicants subject to random assignment includes everyone with marginal priority at lottery schools and applicants with tie-breakers inside the relevant bandwidth at screened schools. Same-color tie-breakers are shared. Schools *1, 3, 5,* and *6* are screened, while *2* and *4* have lottery tie-breakers. The applicant's preferences are $1 \succ_i 2 \succ_i 3 \succ_i 4 \succ_i 5 \succ_i 6$. Arrows mark $\pi_{is} = \rho_{is} + R_{iv(s)}$, the applicant's position at each school s. Lower π_{is} is better. Integers indicate priorities ρ_s, and tick marks indicate the DA cutoff, $\xi_s = \rho_s + \tau_s$. Note that $t_6 = a$, so this applicant is sure to be seated somewhere. The assignment probability therefore sums to 1: if $\tau_2 \geq \tau_4$, the probability of any assignment is $\tau_2 + 0.5 \times (1 - \tau_2) + 0 + 2 \times 0.5^2 \times (1 - \tau_2) = 1$; if $\tau_2 < \tau_4$, this probability is $\tau_2 + 0.5 \times (1 - \tau_2) + 0.5 \times (\tau_4 - \tau_2) + 2 \times 0.5^2 \times (1 - \tau_4) = 1$.

The applicant is never seated at school *1* since his position is to the right of the δ-neighborhood, conditionally seated at schools *2* and *4* since his priority is equal to the marginal priority at each school, conditionally seated at schools *3* and *5* since his position is within the δ-neighborhood at each school, and always seated at school *6* since his position is to the left of the δ-neighborhood.

The columns next to the lines record tie-breaker cutoff, τ, disqualification probability at lottery schools, λ, schools contributing to λ, the disqualification probability at screened schools, σ, schools contributing to σ, and assignment probability.

The local score at each school is computed as follows:

School *1*: The local score at school *1* is zero because $t_{i1}(\delta) = n$.

School *2*: MID at school *2* is zero because this applicant ranks no other lottery school higher. Hence, the second line of (9) applies and probability is given by the tie-breaker cutoff at school *2*, which is τ_2.

School *3*: Since $t_{i3}(\delta) = c$, the third line of (9) applies. The local score at school *3* is the probability of not being assigned to school *2*, that is, $1 - \tau_2$, times 0.5. This last term is the probability associated with being local to the cutoff at school *3*.

School *4*: MID at school *4* is determined by the tie-breaker cutoff at school *2*. When MID exceeds the tie-breaker cutoff at school *4*, then school *4* assignment probability is zero. Otherwise, since $t_{i3}(\delta) = c$ and school *4* is a lottery school, the second line of (9) applies. The probability is therefore 0.5 times the difference between the cutoff at school *4* and MID.

School *5*: MID at school *5* is determined by the larger of the tie-breaker cutoffs at school *2* and school *4*. Since $t_{i5}(\delta) = c$, the third line of (9) applies, and the probability is determined by $(0.5)^2$ times λ, the disqualification probability at lottery schools.

School *6*: Finally, since $t_{i6}(\delta) = a$, the first line of (9) applies and the local score becomes $(0.5)^2$ times λ.

Since $t_{i6}(\delta) = a$, the probabilities sum to 1. If $\tau_2 \geq \tau_4$, the probability of any assignment is $\tau_2 + 0.5 \times (1 - \tau_2) + 2 \times (0.5)^2 \times (1 - \tau_2) = 1$. If $\tau_2 < \tau_4$, the probability is $\tau_2 + 0.5 \times (1 - \tau_4) + 0.5 \times (\tau_4 - \tau_2) + 2 \times 0.5^2 \times (1 - \tau_4) = 1$.

C Additional Results and Proofs

C.1 The DA Propensity Score

This appendix derives the DA propensity score defined as the probability of assignment conditional on type for all applicants, without regard to cutoff proximity. The serial dictatorship propensity score discussed in Section 3.1 is a special case of this.

$MID^v_{\theta s}$ and priority status determine DA propensity score with general tie-breakers. For this proposition, we assume that tie-breakers R_{iv} and $R_{iv'}$ are independent for $v \neq v'$.

Proposition 3 (The DA Propensity Score with General Tie-breaking)**.** *Consider DA with multiple tie-breakers indexed by v, distributed independently of one another according to $F_v(r|\theta)$. For all s and θ in this match,*

$$p_s(\theta) = \begin{cases} 0 & \text{if } \rho_{\theta s} > \rho_s \\ \prod_v (1 - F_v(MID^v_{\theta s}|\theta)) & \text{if } \rho_{\theta s} < \rho_s \\ \prod_{v \neq v(s)} (1 - F_v(MID^v_{\theta s}|\theta)) & \\ \times \max\left\{0, F_{v(s)}(\tau_s|\theta) - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)\right\} & \text{if } \rho_{\theta s} = \rho_s \end{cases}$$

where $F_{v(s)}(\tau_s|\theta) = \tau_s$ and $F_{v(s)}(MID^{v(s)}_{\theta s}|\theta) = MID^{v(s)}_{\theta s}$ when $v(s) \in \{1, ..., U\}$.

Proposition 3, which generalizes an earlier multiple lottery tie-breaker result in Abdulkadiroğlu et al. (2017a), covers three sorts of applicants. First, applicants with less-than-marginal priority at s have no chance of being seated there. The second line of the theorem reflects the likelihood of qualification at schools preferred to s among applicants surely seated at s when they can't do better. Since tie-breakers are assumed independent, the probability of not doing better than s is described by a product over tie-breakers, $\prod_v (1 - F_v(MID^v_{\theta s}|\theta))$. If type θ is sure to do better than s, then $MID^v_{\theta s} = 1$ and the probability at s is zero.

Finally, the probability for applicants with $\rho_{\theta_i s} = \rho_s$ multiplies the term

$$\prod_{v \neq v(s)} (1 - F_v(MID^v_{\theta s}|\theta))$$

by

$$\max\left\{0, F_{v(s)}(\tau_s|\theta) - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)\right\}.$$

The first of these is the probability of failing to improve on s by virtue of being seated at schools using a tie-breaker *other* than $v(s)$. The second parallels assignment probability in single-tie-breaker serial dictatorship: to be seated at s, applicants in $\rho_{\theta_i s} = \rho_s$ must have $R_{iv(s)}$ between $MID^{v(s)}_{\theta s}$ and τ_s.

Proposition 3 allows for single tie-breaking, lottery tie-breaking, or a mix of non-lottery and lottery tie-breakers as in the NYC high school match. With a single tie-breaker, the propensity score formula simplifies, omitting product terms over v:

Corollary 2 (Abdulkadiroğlu et al. (2017a))**.** *Consider DA using a single tie-breaker, R_i, distributed according to $F_R(r|\theta)$ for type θ. For all s and θ in this market, we have:*

$$p_s(\theta) = \begin{cases} 0 & \text{if } \rho_{\theta s} > \rho_s, \\ 1 - F_R(MID_{\theta s}|\theta) & \text{if } \rho_{\theta s} < \rho_s, \\ (1 - F_R(MID_{\theta s}|\theta)) \times \max\left\{0, \dfrac{F_R(\tau_s|\theta) - F_R(MID_{\theta s}|\theta)}{1 - F_R(MID_{\theta s}|\theta)}\right\} & \text{if } \rho_{\theta s} = \rho_s, \end{cases}$$

where $p_s(\theta) = 0$ when $MID_{\theta s} = 1$ and $\rho_{\theta s} = \rho_s$, and $MID_{\theta s}$ is as defined in Section 3, applied to a single tie-breaker.

Common lottery tie-breaking for all schools further simplifies the DA propensity score. When $v(s) = 1$ for all s, $F_R(MID_{\theta s}) = MID_{\theta s}$ and $F_R(\tau_s|\theta) = \tau_s$, as in the Denver match analyzed by Abdulkadiroğlu et al. (2017a). In this case, the DA propensity score is a function only of $MID_{\theta s}$ and the classification of applicants into being never, always, and conditionally seated. This contrasts with the scores in Propositions 2 and 3, which depend on the unknown and unrestricted conditional distributions of tie-breakers given type ($F_R(\tau_s|\theta)$ and $F_R(MID_{\theta s}|\theta)$ with a single tie-breaker; $F_v(\tau_s|\theta)$ and $F_v(MID_{\theta s}|\theta)$ with general tie-breakers). We therefore turn again to the local propensity score to isolate assignment variation that is independent of type and potential outcomes.

Proof of Proposition 3

We prove Proposition 3 using a strategy to that used in the proof of Theorem 1 in Abdulkadiroğlu et al. (2017a). Note first that admissions cutoffs ξ in a large market do not depend on the realized tie-breakers r_{iv}'s: DA in the large market depends on the r_{iv}'s only through $G(I_0)$, defined as the fraction of applicants in set $I_0 = \{i \in I \mid \theta_i \in \Theta_0, r_{iv} \leq r_v \text{ for all } v\}$ with various choices of Θ_0 and r_v. In particular, $G(I_0)$ doesn't depend on tie-breaker realizations in the large market. For the empirical CDF of each tie-breaker conditional on each type, $\hat{F}_v(\cdot|\theta)$, the Glivenko-Cantelli theorem for independent but non-identically distributed random variables implies $\hat{F}_v(\cdot|\theta) = F_v(\cdot|\theta)$ for any v and θ (Wellner, 1981). Since cutoffs ξ are constant, marginal priority ρ_s is also constant for every school s.

Now, consider the propensity score for school s. First, applicants who don't rank s have

$p_s(\theta) = 0$. If $\rho_{\theta s} > \rho_s$, then $\rho_{\theta s} > \rho_s$. Therefore,

$$p_s(\theta) = 0 \text{ if } \rho_{\theta s} > \rho_s \text{ or } \theta \text{ does not rank } s.$$

Second, if $\rho_{\theta s} \leq \rho_s$, then the type θ applicant may be assigned a preferred school $\tilde{s} \in B_{\theta s}$, where $\rho_{\theta \tilde{s}} = \rho_{\tilde{s}}$. For each tie-breaker v, the proportion of type θ applicants assigned some $\tilde{s} \in B^v_{\theta s}$ where $\rho_{\theta \tilde{s}} = \rho_{\tilde{s}}$ is $F_v(MID^v_{\theta s}|\theta)$. This means that for each v, the probability of not being assigned any $\tilde{s} \in B^v_{\theta s}$ where $\rho_{\theta \tilde{s}} = \rho_{\tilde{s}}$ is $1 - F_v(MID^v_{\theta s}|\theta)$. Since tie-breakers are assumed to be distributed independently of one another, the probability of not being assigned any $\tilde{s} \in B_{\theta s}$ where $\rho_{\theta \tilde{s}} = \rho_{\tilde{s}}$ for a type θ applicant is $\Pi_v(1 - F_v(MID^v_{\theta s}|\theta))$. Every applicant of type $\rho_{\theta s} < \rho_s$ who is not assigned a preferred choice is assigned s because $\rho_{\theta s} < \rho_s$. So

$$p_s(\theta) = \Pi_v(1 - F_v(MID^v_{\theta s}|\theta)) \text{ if } \rho_{\theta s} < \rho_s.$$

Finally, consider applicants of type $\rho_{\theta s} = \rho_s$ who are not assigned a choice preferred to s. The fraction of applicants $\rho_{\theta s} = \rho_s$ who are not assigned a preferred choice is $\Pi_v(1 - F_v(MID^v_{\theta s}|\theta))$. Also, the values of the tie-breaking variable $v(s)$ of these applicants are larger than $MID^{v(s)}_{\theta s}$. If $\tau_s < MID^{v(s)}_{\theta s}$, then no such applicant is assigned s. If $\tau_s \geq MID^{v(s)}_{\theta s}$, then the fraction of applicants who are assigned s within this set is given by $\frac{F_{v(s)}(\tau_s|\theta) - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)}{1 - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)}$. Hence, conditional on $\rho_{\theta s} = \rho_s$ and not being assigned a choice higher than s, the probability of being assigned s is given by $\max\{0, \frac{F_{v(s)}(\tau_s|\theta) - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)}{1 - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)}\}$. Therefore,

$$p_s(\theta) = \prod_{v \neq v(s)} (1 - F_v(MID^v_{\theta s}|\theta)) \times \max\left\{0, F_{v(s)}(\tau_s|\theta) - F_{v(s)}(MID^{v(s)}_{\theta s}|\theta)\right\} \text{ if } \rho_{\theta s} = \rho_s.$$

C.2 Proof of Theorem 2

The proof uses lemmas established below. The first lemma shows that the vector of DA cutoffs computed for the sampled market, $\hat{\xi}_N$, converges to the vector of cutoffs in the continuum.

Lemma 3. *(Cutoff almost sure convergence)* $\hat{\xi}_N \xrightarrow{a.s.} \xi$ *where* ξ *denotes the vector of continuum market cutoffs.*

This result implies that the estimated score converges to the large-market local score as market size grows and bandwidth shrinks.

Lemma 4. *(Estimated local propensity score almost sure convergence) For all* $\theta \in \Theta, s \in S$, *and* $T \in \{a, c, n\}^S$, *we have* $\hat{\psi}_s(\theta, T(\delta_N)) \xrightarrow{a.s.} \psi_s(\theta, T)$ *as* $N \to \infty$ *and* $\delta_N \to 0$.

The next lemma shows that the true finite market score with a fixed bandwidth, defined as $\psi_{Ns}(\theta, T; \delta_N) \equiv E_N[D_i(s)|\theta_i = \theta, T_i(\delta_N) = T]$, also converges to $\psi_s(\theta, T)$ as market size grows and bandwidth shrinks.

Lemma 5. *(Bandwidth-specific propensity score almost sure convergence) For all $\theta \in \Theta, s \in S$, $T \in \{a, c, n\}^S$, and δ_N such that $\delta_N \to 0$ and $N\delta_N \to \infty$ as $N \to \infty$, we have $\psi_{Ns}(\theta, T; \delta_N) \xrightarrow{p} \psi_s(\theta, T)$ as $N \to \infty$.*

Finally, the definitions of $\psi_{Ns}(\theta, T; \delta_N)$ and $\psi_{Ns}(\theta, T)$ imply that $|\psi_{Ns}(\theta, T; \delta_N) - \psi_{Ns}(\theta, T)| \xrightarrow{a.s.} 0$ as $\delta_N \to 0$. Combining these results shows that for all $\theta \in \Theta, s \in S$, and T, as $N \to \infty$ and $\delta_N \to 0$ with $N\delta_N \to \infty$, we have

$$
\begin{aligned}
&|\hat{\psi}_s(\theta, T(\delta_N)) - \psi_{Ns}(\theta, T)| \\
=&|\hat{\psi}_s(\theta, T(\delta_N)) - \psi_{Ns}(\theta, T; \delta_N) + \psi_{Ns}(\theta, T; \delta_N) - \psi_{Ns}(\theta, T)| \\
\leq&|\hat{\psi}_s(\theta, T(\delta_N)) - \psi_{Ns}(\theta, T; \delta_N)| + |\psi_{Ns}(\theta, T; \delta_N) - \psi_{Ns}(\theta, T)| \\
\xrightarrow{p}&|\psi_s(\theta, T) - \psi_s(\theta, T)| + 0 \\
=&0.
\end{aligned}
$$

This yields the theorem since Θ, S, and $\{n, c, a\}^S$ are finite.

Proof of Lemma 3

The proof of Lemma 3 is analogous to the proof of Lemma 3 in Abdulkadiroğlu et al. (2017a) and available upon request. The main difference is that to deal with multiple non-lottery tie-breakers, the proof of Lemma 3 needs to invoke the continuous differentiability of $F_v^i(r|e)$ and the Glivenko-Cantelli theorem for independent but non-identically distributed random variables (Wellner, 1981).

Proof of Lemma 4

$\hat{\psi}_s(\theta, T(\delta_N))$ is almost everywhere continuous in finite sample cutoffs $\hat{\xi}_N$, finite sample MIDs ($MID^v_{\theta s}$), and bandwidth δ_N. Since every $MID^v_{\theta s}$ is almost everywhere continuous in finite sample cutoffs $\hat{\xi}_N$, $\hat{\psi}_s(\theta, T(\delta_N))$ is almost everywhere continuous in finite sample cutoffs $\hat{\xi}_N$ and bandwidth δ_N. Recall $\delta_N \to 0$ by assumption while $\hat{\xi}_N \xrightarrow{a.s.} \xi$ by Lemma 3. Therefore, by the continuous mapping theorem, as $N \to \infty$, $\hat{\psi}_s(\theta, T(\delta_N))$ almost surely converges to $\hat{\psi}_s(\theta, T(\delta_N))$ with ξ replacing $\hat{\xi}_N$, which converges to $\psi_s(\theta, T)$ as $\delta_N \to 0$.

Proof of Lemma 5

We use the following fact, which is implied by Example 19.29 in van der Vaart (2000).

Lemma 6. *Let X be a random variable distributed according to some CDF F over $[0,1]$. Let $F(\cdot|X \in [x-\delta, x+\delta])$ be the conditional version of F conditional on X being in a small window $[x-\delta, x+\delta]$ where $x \in [0,1]$ and $\delta \in (0,1]$. Let $X_1, ..., X_N$ be iid draws from F. Let $\hat{F}_N$ be the empirical CDF of $X_1, ..., X_N$. Let $\hat{F}_N(\cdot|X \in [x-\delta, x+\delta])$ be the conditional version of $\hat{F}_N$ conditional on a subset of draws falling in $[x-\delta, x+\delta]$, i.e., $\{X_i | i = 1, ..., n, X_i \in [x-\delta, x+\delta]\}$. Suppose (δ_N) is a sequence with $\delta_N \downarrow 0$ and $\delta_N \times N \to \infty$. Then $\hat{F}_N(\cdot|X \in [x-\delta_N, x+\delta_N])$ uniformly converges to $F(\cdot|X \in [x-\delta_N, x+\delta_N])$, i.e.,*

$$\sup_{x' \in [0,1]} |\hat{F}_N(x'|X \in [x-\delta_N, x+\delta_N]) - F(x'|X \in [x-\delta_N, x+\delta_N])| \to_p 0 \text{ as } N \to \infty \text{ and } \delta_N \to 0.$$

Proof of Lemma 6. We first prove the statement for $x \in (0,1)$. Let P be the probability measure of X and $\hat{P}_N$ be the empirical measure of $X_1, ..., X_N$. Note that

$$\begin{aligned}
&\sup_{x' \in [0,1]} |\hat{F}_N(x'|X \in [x-\delta_N, x+\delta_N]) - F(x'|X \in [x-\delta_N, x+\delta_N])| \\
=& \sup_{t \in [-1,1]} |\hat{F}_N(x+t\delta_N|X \in [x-\delta_N, x+\delta_N]) - F(x+t\delta_N|X \in [x-\delta_N, x+\delta_N])| \\
=& \sup_{t \in [-1,1]} |\frac{\hat{P}_N[x-\delta_N, x+t\delta_N]}{\hat{P}_N[x-\delta_N, x+\delta_N]} - \frac{P_X[x-\delta_N, x+t\delta_N]}{P_X[x-\delta_N, x+\delta_N]}| \\
=& \frac{1}{\hat{P}_N[x-\delta_N, x+\delta_N] P_X[x-\delta_N, x+\delta_N]} \\
&\times \sup_{t \in [-1,1]} |\hat{P}_N[x-\delta_N, x+t\delta_N] P_X[x-\delta_N, x+\delta_N] - \hat{P}_N[x-\delta_N, x+\delta_N] P_X[x-\delta_N, x+t\delta_N]| \\
=& \frac{1}{\hat{P}_N[x-\delta_N, x+\delta_N] P_X[x-\delta_N, x+\delta_N]} \\
&\times \sup_{t \in [-1,1]} |\hat{P}_N[x-\delta_N, x+t\delta_N](P_X[x-\delta_N, x+\delta_N] - \hat{P}_N[x-\delta_N, x+\delta_N]) \\
&\quad + \hat{P}_N[x-\delta_N, x+\delta_N](\hat{P}_N[x-\delta_N, x+t\delta_N] - P_X[x-\delta_N, x+t\delta_N])| \\
\leq& \frac{1}{\hat{P}_N[x-\delta_N, x+\delta_N] P_X[x-\delta_N, x+\delta_N]} \\
&\times \{\sup_{t \in [-1,1]} \hat{P}_N[x-\delta_N, x+t\delta_N] |\hat{P}_N[x-\delta_N, x+\delta_N] - P_X[x-\delta_N, x+\delta_N]| \\
&\quad + \sup_{t \in [-1,1]} \hat{P}_N[x-\delta_N, x+\delta_N] |\hat{P}_N[x-\delta_N, x+t\delta_N] - P_X[x-\delta_N, x+t\delta_N]|\} \\
&= \frac{1}{P_X[x-\delta_N, x+\delta_N]} \times \{|\hat{P}_N[x-\delta_N, x+\delta_N] - P_X[x-\delta_N, x+\delta_N]|
\end{aligned}$$

$$+ \sup_{t\in[-1,1]} |\hat{P}_N[x-\delta_N, x+t\delta_N] - P_X[x-\delta_N, x+t\delta_N]|\}$$
$$= \frac{A_N}{P_X[x-\delta_N, x+\delta_N]},$$

where

$$A_N = |\hat{P}_N[x-\delta_N, x+\delta_N] - P_X[x-\delta_N, x+\delta_N]| + \sup_{t\in[-1,1]} |\hat{P}_N[x-\delta_N, x+t\delta_N] - P_X[x-\delta_N, x+t\delta_N]|.$$

The above inequality holds by the triangle inequality and the second last equality holds because $\sup_{t\in[-1,1]} \hat{P}_N[x-\delta_N, x+t\delta_N] = \hat{P}_N[x-\delta_N, x+\delta_N]$.

We show that $A_N/P_X[x-\delta_N, x+\delta_N] \xrightarrow{p} 0$. Example 19.29 in van der Vaart (2000) implies that the sequence of processes $\{\sqrt{n/\delta_N}(\hat{P}_N[x-\delta_N, x+t\delta_N] - P_X[x-\delta_N, x+t\delta_N]) : t \in [-1,1]\}$ converges in distribution to a Gaussian process in the space of bounded functions on $[-1,1]$ as $N \to \infty$. We denote this Gaussian process by $\{\mathbb{G}_t : t \in [-1,1]\}$. We then use the continuous mapping theorem to obtain

$$\sqrt{n/\delta_N} A_N \xrightarrow{d} |\mathbb{G}_1| + \sup_{t\in[-1,1]} |\mathbb{G}_t|$$

as $N \to \infty$. Since $\{\mathbb{G}_t : t \in [-1,1]\}$ has bounded sample paths, it follows that $|\mathbb{G}_1| < \infty$ and $\sup_{t\in[-1,1]} |\mathbb{G}_t| < \infty$ for sure. By the continuous mapping theorem, under the condition that $N\delta_N \to \infty$,

$$\begin{aligned}(1/\delta_N)A_N &= (1/\sqrt{N\delta_N}) \times \sqrt{n/\delta_N} A_N \\ &\xrightarrow{d} 0 \times (|\mathbb{G}_1| + \sup_{t\in[-1,1]} |\mathbb{G}_t|) \\ &= 0.\end{aligned}$$

This implies that $(1/\delta_N)A_N \xrightarrow{p} 0$, because for any $\epsilon > 0$,

$$\begin{aligned}\Pr(|(1/\delta_N)A_N| > \epsilon) &= \Pr((1/\delta_N)A_N < -\epsilon) + \Pr((1/\delta_N)A_N > \epsilon) \\ &\leq \Pr((1/\delta_N)A_N \leq -\epsilon) + 1 - \Pr((1/\delta_N)A_N \leq \epsilon) \\ &\to \Pr(0 \leq -\epsilon) + 1 - \Pr(0 \leq \epsilon) \\ &= 0,\end{aligned}$$

where the convergence holds since $(1/\delta_N)A_N \xrightarrow{d} 0$. To show that $A_N/P_X[x-\delta_N, x+\delta_N] \xrightarrow{p}$

0, it is therefore enough to show that $\lim_{N\to\infty}(1/\delta_N)P_X[x-\delta_N, x+\delta_N] > 0$. We have

$$\begin{aligned}(1/\delta_N)P_X[x-\delta_N, x+\delta_N] &= (1/\delta_N)(F_X(x+\delta_N) - F_X(x-\delta_N)) \\ &= (1/\delta_N)(2f(x)\delta_N + o(\delta_N)) \\ &= 2f(x) + o(1) \\ &\to 2f(x) \\ &> 0,\end{aligned}$$

where we use Taylor's theorem for the second equality and the assumption of $f(x) > 0$ for the last inequality.

We next prove the statement for $x = 0$. Note that

$$\begin{aligned}&\sup_{x'\in[0,1]} |\hat{F}_N(x'|X \in [-\delta_N, \delta_N]) - F(x'|X \in [-\delta_N, \delta_N])| \\ =& \sup_{t\in[0,1]} |\hat{F}_N(t\delta_N|X \in [0, \delta_N]) - F(t\delta_N|X \in [0, \delta_N])| \\ =& \sup_{t\in[0,1]} |\frac{\hat{F}_N(t\delta_N)}{\hat{F}_N(\delta_N)} - \frac{F_X(t\delta_N)}{F_X(\delta_N)}| \\ =& \frac{1}{\hat{F}_N(\delta_N)F_X(\delta_N)} \sup_{t\in[0,1]} |\hat{F}_N(t\delta_N)F_X(\delta_N) - \hat{F}_N(\delta_N)F_X(t\delta_N)| \\ =& \frac{1}{\hat{F}_N(\delta_N)F_X(\delta_N)} \sup_{t\in[0,1]} |\hat{F}_N(t\delta_N)(F_X(\delta_N) - \hat{F}_N(\delta_N)) + \hat{F}_N(\delta_N)(\hat{F}_N(t\delta_N) - F_X(t\delta_N))| \\ \le& \frac{1}{\hat{F}_N(\delta_N)F_X(\delta_N)}\{\sup_{t\in[0,1]} \hat{F}_N(t\delta_N)|\hat{F}_N(\delta_N) - F_X(\delta_N)| + \sup_{t\in[0,1]} \hat{F}_N(\delta_N)|\hat{F}_N(t\delta_N) - F_X(t\delta_N)|\} \\ =& \frac{1}{F_X(\delta_N)}\{|\hat{F}_N(\delta_N) - F_X(\delta_N)| + \sup_{t\in[0,1]} |\hat{F}_N(t\delta_N) - F_X(t\delta_N)|\} = \frac{A_N^0}{F_X(\delta_N)},\end{aligned}$$

where $A_N^0 = |\hat{F}_N(\delta_N) - F_X(\delta_N)| + \sup_{t\in[0,1]} |\hat{F}_N(t\delta_N) - F_X(t\delta_N)|$. By the argument used in the above proof for $x \in (0, 1)$, we have $(1/\delta_N)A_N^0 \xrightarrow{p} 0$. It also follows that

$$\begin{aligned}(1/\delta_N)F_X(\delta_N) &= (1/\delta_N)(f(0)\delta_N + o(\delta_N)) \\ &= f(0) + o(1) \\ &\to f(0) \\ &> 0.\end{aligned}$$

Thus, $\frac{A_N^0}{F_X(\delta_N)} \xrightarrow{p} 0$, and hence $\sup_{x'\in[0,1]} |\hat{F}_N(x'|X \in [-\delta_N, \delta_N]) - F(x'|X \in [-\delta_N, \delta_N])| \xrightarrow{p} 0$. The proof for $x = 1$ follows from the same argument. □

Consider any deterministic sequence of economies $\{g_N\}$ such that $g_N \in \mathcal{G}$ for all N and $g_N \to G$ in the $(\mathcal{G}, d)$ metric space. Let (δ_N) be an associated sequence of positive numbers (bandwidths) such that $\delta_N \to 0$ and $N\delta_N \to \infty$ as $N \to \infty$. Let $\psi_{Ns}(\theta, T; \delta_N) \equiv E_N[D_i(s)|\theta_i = \theta, T_i(\delta_N) = T]$ be the (finite-market, deterministic) bandwidth-specific propensity score for particular g_N and δ_N.

For Lemma 5, it is enough to show deterministic convergence of this finite-market score, that is, $\psi_{Ns}(\theta, T; \delta_N) \to \psi_s(\theta, T)$ as $g_N \to G$ and $\delta_N \to 0$. To see this, let G_N be the distribution over $I(\Theta_0, r_0, r_1)$'s induced by randomly drawing N applicants from G, where $I(\Theta_0, r_0, r_1) \equiv \{i|\theta_i \in \Theta_0, r_0 < r_i \leq r_1\}$. Note that G_N is random and that $G_N \overset{a.s.}{\to} G$ by Wellner (1981)'s Glivenko-Cantelli theorem for independent but non-identically distributed random variables. $G_N \overset{p}{\to} G$ and $\psi_{Ns}(\theta, T; \delta_N) \to \psi_s(\theta, T)$ allow us to apply the Extended Continuous Mapping Theorem (Theorem 18.11 in van der Vaart (2000)) to obtain $\tilde{\psi}_{Ns}(\theta, T; \delta_N) \overset{p}{\longrightarrow} \psi_s(\theta, T)$ where $\tilde{\psi}_{Ns}(\theta, T; \delta_N)$ is the random version of $\psi_{Ns}(\theta, T; \delta_N)$ defined for G_N.

For notational simplicity, consider the single-school RD case, where there is only one school s making assignments based on a single non-lottery tie-breaker $v(s)$ (without using any priority). A similar argument with additional notation shows the result for DA with general tie-breaking.

For any $\delta_N > 0$, whenever $T_i(\delta_N) = a$, it is the case that $D_i(s) = 1$. As a result,

$$\psi_{Ns}(\theta, a; \delta_N) \equiv E_N[D_i(s)|\theta_i = \theta, T_i(\delta_N) = a] = 1 \equiv \psi_s(\theta, a).$$

Therefore, $\psi_{Ns}(\theta, a; \delta_N) \to \psi_s(\theta, a)$ as $N \to \infty$. Similarly, for any $\delta_N > 0$, whenever $T_i(\delta_N) = n$, it is the case that $D_i(s) = 0$. As a result,

$$\psi_{Ns}(\theta, n; \delta_N) \equiv E_N[D_i(s)|\theta_i = \theta, T_i(\delta_N) = n] = 0 \equiv \psi_s(\theta, n).$$

Therefore, $\psi_{Ns}(\theta, n; \delta_N) \to \psi_s(\theta, n)$ as $N \to \infty$. Finally, when $T_i(\delta_N) = c$, let

$$F_{N,v(s)}(r|\theta) \equiv \frac{\sum_{i=1}^N 1\{\theta_i = \theta\} F^i_{v(s)}(r)}{\sum_{i=1}^N 1\{\theta_i = \theta\}}$$

be the aggregate tie-breaker distribution conditional on each applicant type θ in the finite market. $\tilde{\xi}_{Ns}$ denotes the random cutoff at school s in a realized economy g_N. For any ϵ, there exists N_0 such that for any $N > N_0$, we have

$$\psi_{Ns}(\theta, c; \delta_N) \equiv E_N[D_i(s)|\theta_i = \theta, T_i(\delta_N) = c]$$

$$
\begin{aligned}
&= P_N[R_{iv(s)} \le \tilde{\xi}_{Ns} | \theta_i = \theta, R_{iv(s)} \in (\tilde{\xi}_{Ns} - \delta_N, \tilde{\xi}_{Ns} + \delta_N]] \\
&\in (P[R_{iv(s)} \le \xi_s | \theta_i = \theta, R_{iv(s)} \in (\xi_s - \delta_N, \xi_s + \delta_N]] - \epsilon/2, \\
&\quad P[R_{iv(s)} \le \xi_s | \theta_i = \theta, R_{iv(s)} \in (\xi_s - \delta_N, \xi_s + \delta_N]] + \epsilon/2),
\end{aligned}
$$

where ξ_s is school s's continuum cutoff, P is the probability induced by the tie-breaker distributions in the continuum economy, and the inclusion is by Assumption 2 and Lemmata 3 and 6. Again for any ϵ, there exists N_0 such that for any $N > N_0$, we have

$$
\begin{aligned}
&(P[R_{iv(s)} \le \xi_s | \theta_i = \theta, R_{iv(s)} \in (\xi_s - \delta_N, \xi_s + \delta_N]] - \epsilon/2, \\
&\ P[R_{iv(s)} \le \xi_s | \theta_i = \theta, R_{iv(s)} \in (\xi_s - \delta_N, \xi_s + \delta_N]] + \epsilon/2) \\
&= (\frac{F_{v(s)}(\xi_s|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)}{F_{v(s)}(\xi_s + \delta_N|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)} - \epsilon/2, \\
&\quad \frac{F_{v(s)}(\xi_s|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)}{F_{v(s)}(\xi_s + \delta_N|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)} + \epsilon/2) \\
&= (\frac{\{F_{v(s)}(\xi_s|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)\}/\delta_N}{\{F_{v(s)}(\xi_s + \delta_N|\theta) - F_{v(s)}(\xi_s|\theta)\}/\delta_N + \{F_{v(s)}(\xi_s|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)\}/\delta_N} - \epsilon/2, \\
&\quad \frac{\{F_{v(s)}(\xi_s|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)\}/\delta_N}{\{F_{v(s)}(\xi_s + \delta_N|\theta) - F_{v(s)}(\xi_s|\theta)\}/\delta_N + \{F_{v(s)}(\xi_s|\theta) - F_{v(s)}(\xi_s - \delta_N|\theta)\}/\delta_N} + \epsilon/2) \\
&\in (0.5 - \epsilon, 0.5 + \epsilon) \\
&= (\psi_s(\theta, c) - \epsilon, \psi_s(\theta, c) + \epsilon),
\end{aligned}
$$

completing the proof.

D Empirical Appendix

D.1 Data

The NYC DOE provided data on students, schools, the rank-order lists submitted by match participants, school assignments, and outcome variables. Applicants and programs are uniquely identified by a number that can be used to merge data sets. Students with a record in assignment files who cannot be matched to other files are omitted.

D.1.1 Applicant Data

We focus on first-time applicants to the NYC public (unspecialized) high school system who live in NYC and attended a public middle school in eighth grade. The NYC high school match is conducted in three rounds. The data used for the present analyses are from the first assignment round, which uses DA and we refer to as *main round*. Applicants who were not assigned after the main round apply to the remaining seats in a subsequent *supplementary round*. Students who remain unassigned in the supplementary round are then assigned on a case-by-case basis in the final *administrative round*.

Assignment, Priorities, and Ranks

Data on the assignment system come from the DOE's enrollment office, and report assignments for our two cohorts. The main application data set details applicant program choices, eligibility, priority group and rank, as well as the admission procedure used at the respective program. Lottery numbers and details on assignments at Educational Option (Ed-Opt) programs are provided in separate data sets.

Student Characteristics

NYC DOE students files record grade, gender, ethnicity, and whether students attended a public middle school. Separate files include (i) student scores on middle school standardized tests, (ii) English language learner and special education status, and (iii) subsidized lunch status. Our baseline middle school scores are from 6th grade math and English exams. If a student re-took a test, the latest result is used. Our demographic characteristics come from the DOE's snapshot for 8th grade.

D.1.2 School-level Data

School Letter Grades

School grades are drawn from NYC DOE School Report Cards for 2010/11, 2011/12 and 2012/13. For each application cohort, we grade schools based on the report cards published in the school year prior to the application school year: for the 2011/12 application cohort, for instance, schools are assigned grades published in 2010/11, and similarly for the other two cohorts.

School Characteristics

School characteristics were taken from report card files provided by the DOE. These data provide information on enrollment statistics, racial composition, attendance rates, suspensions, teacher numbers and experience, and graduating class Regents Math and English performance. A unique identifier for each school allows these data to be merged with data from other sources. The analyses on teacher experience and education reported in Table 2 of this publication are based on the School-Level Master File 1996-2016, a dataset compiled by the Research Alliance for NYC Schools at New York University' Steinhardt School of Culture, Education, and Human Development (www.ranycs.org). All data in the School-Level Master File are publicly available. The Research Alliance takes no responsibility for potential errors in the dataset or the analysis. The opinions expressed in this publication are those of the authors and do not represent the views of the Research Alliance for NYC Schools or the institutions that posted the original publicly available data.[26]

Defining Screened and Lottery Schools

We define lottery schools as any school hosting at least one program for which the lottery number is used as the tie-breaker. Screened schools are the remaining schools. Some schools allow students to share a screened tie-breaker rank, breaking screening-variable ties with lottery numbers. Propensity scores for such schools are computed using the lottery tie breaker and schools are considered lottery in any analysis that makes this substantive distinction. Specialized high schools are considered screened schools. The remaining schools, mostly charters that conduct a separate lottery process, are considered lottery schools.

D.1.3 SAT and Graduation Outcomes

SAT Tests

The NYC DOE has data on SAT scores for test-takers from 2006-17. These data originate with the College Board. We use the first test for multiple takers. For applicants tested in the same month, we use the highest score. During our sample period, the SAT has been

[26] Research Alliance for New York City Schools (2017). School-Level Master File 1996-2016 [Data file and code book]. Unpublished data

redesigned. We re-scale scores of SAT exams taken prior to the reform according to the official re-scaling scheme provided by CollegeBoard.[27]

Graduation

The DOE Graduation file records the discharge status for public school students enrolled from 2005-17. Because data on graduation results are not yet available for the youngest (2013/14) cohort, graduation results are for the two older cohorts only.

College- and Career-preparedness and College-readiness

The DOE provided us with individual-level indicators for college- and career-preparedness as well as college-readiness for public school students enrolled from 2005-17. Since these data are not yet available for the youngest (2013/14) cohort, the results are for the two older cohorts only. Table D1 gives an overview on the criteria for the two indicators.

Table D1. Criteria for College- and Career-preparedness and College-readiness Indicators

College- and Career-preparedness
Any of the following:
- Scored 65+ on the Algebra II, Math B, Chemistry, or Physics Regents exam
- Scored 3+ on any Advanced Placement (AP) or 4+ on any International Baccalaureate (IB) exam
- Earned "C" or higher in a college credit-bearing course or passed another course certified by the DOE
- Earned a diploma with a Career and Technical Education (CTE) endorsement
- Earned a diploma with an Arts endorsement; or passed an industry-recognized technical assessment
College-readiness
For ELA, any of the following:
- SAT Evidence-Based Reading and Writing (EBRW) section score of 480+
- ACT English score of 20+ or NY State English Regents score of 75+
For Math, any of the following:
- SAT Math Section score of 530+
- ACT Math score of 21+
- Common Core Regents: Score of 70+ in Algebra I or 70+ in Geometry or 65+ in Algebra 2
- Other Regents: Score of 80+ in Integrated Algebra or Geometry or Algebra 2/Trigonometry and successful completion of the Algebra 2/Trigonometry or higher-level course
- Score of 75+ in Regents Math A or Math B, or Sequential II or Sequential III

D.1.4 Replicating the NYC Match

NYC uses the student-proposing DA algorithm to determine assignments. The three ingredients for this algorithm are: student's ranking of up to 12 programs, program capacities and priorities, and tie-breakers.

[27] See `https://collegereadiness.collegeboard.org/educators/higher-ed/scoring/concordance` for the conversion scale.

Program Assignment Rules

Programs use a variety of assignment rules. Lottery, Limited Unscreened, and Zoned programs order students first by priority group, and within priority group by lottery number. Screened and Audition programs order students by priority group and then by a non-lottery tie-breaker, referred to as running or rank variable. We observe these in the form of an an ordering of applicants provided by Screened and Audition programs. Ed-Opt programs use two tie-breakers, which is described into more detail below. Finally, as mentioned above, some schools allow students to share a screened tie-breaker rank, breaking screening-variables ties with lottery numbers.

Program Capacities and Priorities

Program capacities must be imputed. We assume program capacity equals the number of assignments extended. Program type determines priorities. The priority group is a number assigned by the NYC DOE depending on addresses, program location, siblings, among other considerations, including, in some cases, whether applicants attended an information session or open house (for Limited Unscreened programs).

Lottery Numbers

The lottery numbers are provided by the NYC DOE in a separate data set. Lottery tie-breakers are reported as unique alphanumeric string and scaled to $[0, 1]$. Lottery numbers are missing for some; we assign these applicants a randomly drawn lottery number and use it in our replicated match. It is this replicated match that is used to construct assignment instruments and their associated propensity scores.

Ranks

Screened, Audition, and half of the seats at Ed-Opt programs assign students a rank, based on various diverging criteria, such as former test performance. Ranks are reported as an integer reflecting raw tie-breaker order in this group. We scale these so as to lie in $(0, 1]$ by transforming raw tie-breaking realizations R_{iv} into $[R_{iv} - \min_j R_{jv} + 1]/[\max_j R_{jv} - \min_j R_{jv} + 1]$ for each tie-breaker v. At some screened programs, the rank numbers of applicants have gaps, i.e. the distribution of running variable values is discontinuous. Potential reasons include i) human error when school principals submit applicant rankings to the NYC DOE, and ii) while running variables are assigned at the program level, applications at Ed-Opt programs are treated as six separate buckets (i.e. distinct application choices), leading to artificial gaps in rank distributions (see discussion of assignment at Ed-Opt programs below).

Assignment at Educational Option programs

Ed-Opt programs use two tie-breakers. Applicants are first categorized into high performers, middle performers, and low performers by scores on a seventh grade reading test. Ed-Opt programs aim to have an enrollment distribution of 16% high performers, 68% middle performers and 16% low performers. Half of Ed-Opt seats are assigned using the lottery tie-breaker. These seats are called "random." The other half uses a rank variable such as those used by other screened programs. These seats are called "select."

We refer to the resulting six combinations as "buckets." Ed-Opt applicants are treated as applying to all six. A separate data set details which bucket applicants were offered. Buckets have their own priorities and capacities. The latter are imputed based on the observed assignments to buckets.

Tables D2 and D3 show applicants' choice order of and priorities at Ed-Opt buckets, respectively. Both are based on consultations with the NYC DOE and our simulations of the match.

High performers rank high buckets first, while medium and low performers apply to medium and low buckets first, respectively.

Table D2. Applicants' Choice Order of EdOpt Buckets

Choice order	High performers		Middle performers		Low performers	
1	High	Select	Middle	Select	Low	Select
2	High	Random	Middle	Random	Low	Random
3	Middle	Select	High	Select	High	Select
4	Middle	Random	High	Random	High	Random
5	Low	Select	Low	Select	Middle	Select
6	Low	Random	Low	Random	Middle	Random

High performers have highest priority (priority group 1) at high buckets, while medium and low performers receive highest priority at medium and low buckets, respectively.

Table D3. Priorities at EdOpt Buckets

Priority	High	Middle	Low
Group	Random or Select	Random or Select	Random or Select
1	High performers	Middle performers	Low performers
2	Middle performers	Low performers	Middle performers
3	Low performers	High performers	High performers

Miscellaneous Sample Restrictions The analysis sample is limited to first-time eighth

grade applicants for ninth grade seats. Ineligible applications (as indicated in the main application data set) are dropped. Applicants with special education status compete for a different set of seats and are thus dropped in the analysis.
Students in the top 2% of scorers on the seventh grade reading are automatically admitted into any Ed-Opt program they rank first. We gather these assignments in a separate Ed-Opt bucket, thereby leaving the admission process to the other six unaffected.

Table D4 records the proportion of applicants for which our match replication was successful.

Table D4. Replication Rates

	Application Cohort		
	2011/2012	2012/2013	2013/2014
	(1)	(2)	(3)
All schools	0.967	0.959	0.967
Grade A schools	0.971	0.962	0.974
Grade A screened schools	0.990	0.988	0.990
Grade A lottery schools	0.965	0.956	0.965

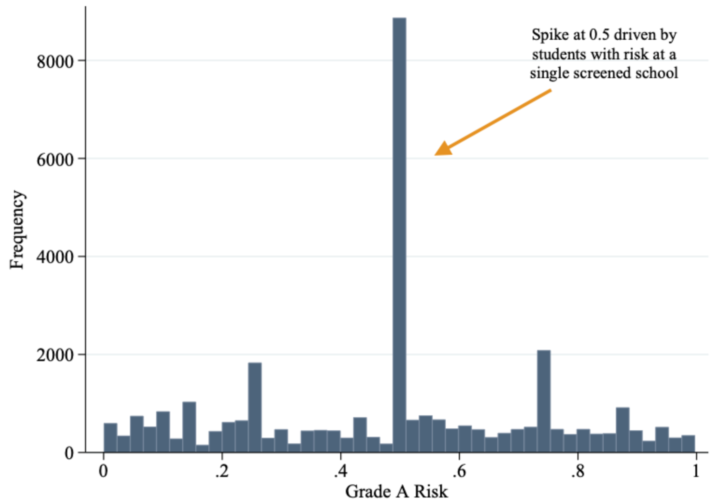
Notes: This table shows replication rates for the New York City match for the three application cohorts in the analysis sample. A replicated offer is one where the offer generated by our run of the match coincides the offer received.

D.2 Additional Empirical Results

Grade A risk has a mode at 0.5, but takes on many other values as well. A probability of 0.5 arises when the overall Grade A propensity score is generated by a single Grade A screened school. This can be seen in Figure D1, which tabulates the estimated probability of assignment to a Grade A school for applicants in all cohorts (2012-2014) with a probability strictly between 0 and 1 calculated using the formula in Theorem 1. There are 26,555 students with the estimated assignment probability equal to 1, 87,742 students with the propensity score equal to 0, and 32,866 students with Grade A risk. The propensity score of 0.5 arises when the overall Grade A propensity score is generated by a single Grade A screened school.

Table D5 reports estimates of the effect of Grade A assignments on attrition, computed by estimating models like those used to gauge balance. Applicants who receive Grade A school assignments have a slightly higher likelihood of taking the SAT. Decomposing Grade A schools into screened and lottery schools, applicants who receive lottery Grade A school

Figure D1: Distribution of Grade A Risk



Notes: This figure shows the histogram of the estimated probability of assignment to a Grade A school for at-risk applicants in all sample cohorts (2012-2014), calculated using Theorem 1. The full sample includes 26,555 applicants with a Grade A propensity score equal to 1, 87,742 applicants with propensity score equal to 0, and 32,866 students with Grade A risk. The at-risk sample is used to compute the balance estimates reported in columns 3 and 4 of Table 3.

assignments are 1.6 percent more likely to have SAT scores, while assignments to Grade A screened schools do not correspond to a statistically significant difference in the likelihood of having follow-up SAT scores. This modest difference seems unlikely to bias the 2SLS Grade A estimates reported in Tables 4 and 5.

Table D6 reports estimates of the effect of enrollment in an ungraded high school. These use models like those used to compute the estimates presented in Table 4. OLS estimates show a small positive effect of ungraded school attendance on SAT scores and a strong negative effect on graduation outcomes. 2SLS estimates, by contrast, suggest ungraded school attendance is unrelated to these outcomes.

Table D5. Differential Attrition

	Non-offered	Grade A School Type		
		Any	Screened	Lottery
	(1)	(2)	(3)	(4)
Took SAT exam	0.765	0.015	-0.004	0.016
		(0.006)	(0.011)	(0.007)
N		32,866	12,002	27,269
Enrolled in ninth grade	0.986	0.005	-0.001	0.006
		(0.002)	(0.003)	(0.002)
N		32,866	12,002	27,269

Notes. This table reports differential attrition estimates, computed by regressing covariates on dummies indicating a Grade A offer and an ungraded school offer, controlling for saturated Grade A and ungraded school propensity scores (columns 2-4), and running variable controls (columns 2 and 3). Screened program bandwidths are calculated as suggested by Imbens and Kalyanaraman (IK; 2012) with a uniform kernel. See the text for details. Robust standard errors are in parenthesis.

Table D6. 2SLS Estimates of the Effect of Attending an Ungraded School

	All applicants		Applicants with Grade A Risk	
	Non-enrolled mean	OLS	Non-offered mean	2SLS
	(1)	(2)	(3)	(4)
SAT Math	474	1.20	517	1.41
(200-800)	(103)	(0.189)	(109)	(1.79)
SAT Reading	474	1.06	512	0.237
(200-800)	(90)	(0.176)	(93)	(1.72)
N		124,902		24,707
Graduated	0.739	-0.236	0.825	0.034
		(0.003)		(0.025)
N		183,526		31,976
College- and Career-prepared	0.429	-0.134	0.595	0.034
		(0.003)		(0.037)
College-ready	0.374	-0.096	0.550	0.021
		(0.003)		(0.036)
N		121,416		20,664

Notes. This table reports OLS and 2SLS estimates of ungraded school effects produced by the models reported in columns 3 and 4 of Table 4. Robust standard errors are in parenthesis.